(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁷: **C12N 15/63**, 15/12, A61K 39/395, 48/00, C07K 16/18, 14/435

(21) International Application Number: PCT/AU02/00136

(22) International Filing Date: 12 February 2002 (12.02.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PR 3052   12 February 2001 (12.02.2001)   AU
PR 3053   12 February 2001 (12.02.2001)   AU
PR 3054   12 February 2001 (12.02.2001)   AU

(71) Applicant *(for all designated States except US)*: **BIONOMICS LIMITED** [AU/AU]; 31 Dalgleish Street, Thebarton, S.A. 5031 (AU).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **KREMMIDIOTIS, Gabriel** [AU/AU]; 1 Nina Court, Aberfoyle Park, S.A. 5159 (AU). **BAIS, Anthony, John** [AU/AU]; 20 Broughton Avenue, Kurralta Park, S.A. 5037 (AU). **CALLEN, David, Frederick** [AU/AU]; 40 Clifton Street, Malvern, S.A. 5061 (AU). **WHITMORE, Scott, Anthony**

[AU/AU]; 17 James Harrold Court, Golden Grove, S.A. 5125 (AU).

(74) **Agent: NUGENT, Brendan, John**; Griffith Hack, 167 Eagle Street, Brisbane, QLD 4000 (AU).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DNA SEQUENCES DIFFERENTIALLY EXPRESSED IN TUMOUR CELL LINES

(57) **Abstract:** A method for the diagnosis of cancer, or a predisposition thereto, in a patient, comprising the steps of: (1) establishing the level of expression of a gene selected from the group consisting of cytochrome B-245, alpha polypeptide (CYBA), cadherin 15 (CDH15), dipeptidase 1 (DPEP1), solute carrier family 7, member 5 (SLC7A5), BNO8, BNO205 and BNO221; and (2) comparing expression of the gene to a baseline established from expression in normal tissue controls; wherein substantial variants from the baseline indicates that the patient is susceptible to cancer.

DNA SEQUENCES DIFFERENTIALLY EXPRESSED IN TUMOUR CELL LINES

Technical Field

The present invention is concerned with DNA sequences
from the 16q24.3 region which have been shown to be
differentially expressed in breast cancer cell lines, and
are therefore implicated in the development of carcinomas.

Background Art

The development of human carcinomas has been shown to
arise from the accumulation of genetic changes involving
both positive regulators of cell function (oncogenes) and
negative regulators (tumour suppressor genes). For a
normal somatic cell to evolve into a metastatic tumour it
requires changes at the cellular level, such as
immortalisation, loss of contact inhibition and invasive
growth capacity, and changes at the tissue level, such as
evasion of host immune responses and growth restraints
imposed by surrounding cells, and the formation of a blood
supply for the growing tumour.

Molecular genetic studies of colorectal carcinoma
have provided substantial evidence that the generation of
malignancy requires the sequential accumulation of a
number of genetic changes within the same epithelial stem
cell of the colon. For a normal colonic epithelial cell to
become a benign adenoma, progress to intermediate and late
adenomas, and finally become a malignant cell,
inactivating mutations in tumour suppressor genes and
activating mutations in proto-oncogenes are required
(Fearon and Vogelstein,1990).

The employment of a number of techniques, such as
loss of heterozygosity (LOH), comparative genomic
hybridisation (CGH) and cytogenetic studies of cancerous
tissue, all of which exploit chromosomal abnormalities
associated with the affected cell, has aided in the
identification of a number of tumour suppressor genes and
oncogenes associated with a range of tumour types.

- 2 -

In one aspect, studies of cancers such as retinoblastoma and colon carcinoma have supported the model that LOH is a specific event in the pathogenesis of cancer and has provided a mechanism in which to identify
5   the cancer causing genes. For instance in colorectal carcinoma, inherited forms of the disease have been mapped to the long arm of chromosome 5 while LOH at 5q has been reported in both the familial and sporadic versions of the disease. The *APC* tumour suppressor gene, mapping to this
10 region, was subsequently shown to be involved (Groden *et al.*, 1991). The model is further highlighted in Von Hippel-Lindau (VHL) syndrome, a rare disorder that predisposes individuals to a variety of tumours including clear cell carcinomas of the kidneys and islet cell
15 tumours of the pancreas. Both sporadic and inherited cases of the syndrome show LOH for the short arm of chromosome 3 and somatic translocations involving 3p in sporadic tumours, and genetic linkage to the same region in affected families has also been observed. The VHL tumour
20 suppressor gene has since been identified from this region of chromosome 3 and mutations in it have been detected in 100% of patients who carry a clinical diagnosis of VHL disease. In addition, the VHL gene is inactivated in approximately 50-80% of the more common sporadic form of
25 renal clear cell carcinoma.

The genetic determinants involved in breast cancer are not as well defined as that of colon cancer due in part to the histological stages of breast cancer development being less well characterised. However, as
30 with colon carcinoma, it is believed that a number of genes need to become involved in a stepwise progression during breast tumourigenesis.

Certain women appear to be at an increased risk of developing breast cancer. Genetic linkage analysis has
35 shown that 5 to 10% of all breast cancers are due to at least two autosomal dominant susceptibility genes. Generally, women carrying a mutation in a susceptibility

gene develop breast cancer at a younger age compared to the general population, often have bilateral breast tumours, and are at an increased risk of developing cancers in other organs, particularly carcinoma of the
5 ovary.

Genetic linkage analysis on families showing a high incidence of early-onset breast cancer (before the age of 46) was successful in mapping the first susceptibility gene, BRCA1, to chromosome 17q21 (Hall et al., 1990).
10 Subsequent to this, the BRCA2 gene was mapped to chromosome 13q12-q13 (Wooster et al., 1994) with this gene conferring a higher incidence of male breast cancer and a lower incidence of ovarian cancer when compared to BRCA1.

Both BRCA1 and BRCA2 have since been cloned (Miki et
15 al., 1994; Wooster et al., 1995) and numerous mutations have been identified in these genes in susceptible individuals with familial cases of breast cancer.

Additional inherited breast cancer syndromes exist, however they are rare. Inherited mutations in the TP53
20 gene have been identified in individuals with Li-Fraumeni syndrome, a familial cancer resulting in epithelial neoplasms occurring at multiple sites including the breast. Similarly, germline mutations in the MMAC1/PTEN gene involved in Cowden's disease and the ataxia
25 telangiectasia (AT) gene have been shown to confer an increased risk of developing breast cancer, among other clinical manifestations, but together account for only a small percentage of families with an inherited predisposition to breast cancer.
30 Somatic mutations in the TP53 gene have been shown to occur in a high percentage of individuals with sporadic breast cancer. However, although LOH has been observed at the BRCA1 and BRCA2 loci at a frequency of 30 to 40% in sporadic cases (Cleton-Jansen et al., 1995; Saito et al.,
35 1993), there is virtually no sign of somatic mutations in the retained allele of these two genes in sporadic cancers (Futreal et al., 1994; Miki et al., 1996). Recent data

- 4 -

suggests that DNA methylation of the promoter sequence of these genes may be an important mechanism of down-regulation. The use of both restriction fragment length polymorphisms and small tandem repeat polymorphic markers has identified numerous regions of allelic imbalance in breast cancer suggesting the presence of additional genes, which may be implicated in breast cancer. Data compiled from more than 30 studies reveals the loss of DNA from at least 11 chromosome arms at a frequency of more than 25%, with regions such as 16q and 17p affected in more than 50% of tumours (Devilee and Cornelisse, 1994; Brenner and Aldaz, 1995). However only some of these regions are known to harbour tumour suppressor genes shown to be mutated in individuals with both sporadic (*TP53* and *RB* genes) and familial (*TP53, RB, BRCA1,* and *BRCA2* genes) forms of breast cancer.

    Cytogenetic studies have implicated loss of the long arm of chromosome 16 as an early event in breast carcinogenesis since it is found in tumours with few or no other cytogenetic abnormalities. Alterations in chromosome 1 and 16 have also been seen in several cases of ductal carcinoma *in situ* (DCIS), the preinvasive stage of ductal breast carcinoma. In addition, LOH studies on DCIS samples identified loss of 16q markers in 29 to 89% of the cases tested (Chen et al., 1996; Radford et al., 1995). In addition, examination of tumours from other tissue types have indicated that 16q LOH is also frequently seen in prostate, lung, hepatocellular, ovarian, primitive neuroectodermal and Wilms' tumours.

    Together, these findings suggest the presence of a gene mapping to the long arm of chromosome 16 that is critically involved in the early development of a large proportion of breast cancers as well as cancers from other tissue types, but to date no such gene has been identified.

- 5 -

Disclosure of the Invention

        The present invention provides nucleic acid and
protein sequences that are differentially expressed in
breast cancer when compared to normal tissue controls,
5 here-in termed "breast cancer sequences". As outlined
below, breast cancer sequences that are differentially
expressed include those that are down-regulated in breast
cancer (tumour suppressor genes) as well as those that are
up-regulated   in   breast   cancer   (oncogenes).   The
10 differential expression of these sequences in breast
cancer combined with the fact they have been identified
from a region of LOH seen in breast cancer as well as
other carcinomas including prostate tumours suggests they
are contributory factors in cancer. The breast cancer
15 sequences of the invention are described in Table 1 and
are represented by SEQ ID Numbers: 1-11.

        "Down-regulation" as used herein means at least about
a 15 to 49 fold decrease in expression, preferably at
least about a 50 to 79 fold decrease in expression, with
20 at least about an 80 fold or higher decrease in expression
being preferred (assuming a relative fold variability
index of 50 or higher).

        "Up-regulation" as used herein means at least about a
15 to 49 fold increase in expression, preferably at least
25 about a 50 to 79 fold increase in expression, with at
least about an 80 fold or higher increase in expression
being preferred (assuming a relative fold variability
index of 50 or higher).

        The present invention also encompasses isolated
30 nucleic acid and/or amino acid sequences which are
homologous to the breast cancer sequences described above.
Such homology is based on the overall nucleic acid or
amino acid sequence of the group described in Table 1 and
represented by the SEQ ID Numbers: 1-11_and is determined
35 using either homology programs or hybridisation conditions
as outlined below.

        A nucleic acid or protein is a breast cancer nucleic

- 6 -

acid or protein if the overall homology of the nucleic
acid or protein sequence to one of the sequences described
in Table 1 and represented by the SEQ ID Numbers: 1-11 is
at least 70%, preferably 85% and most preferably 95%.
5   Homology in this context means sequence similarity or
identity, with identity being preferred.

In a preferred embodiment, the sequences which are
used to determine sequence identity or similarity are
selected from the sequences described in Table 1 and
10   represented by the SEQ ID Numbers: 1-11 or are naturally
occurring allelic variants, sequence variants or splice
variants of these sequences.

Sequence identity is typically calculated using the
BLAST algorithm, described in Altschul *et al Nucleic Acids*
15   *Res.* 25, 3389-3402 (1997) with the BLOSUM62 default
matrix.

In one embodiment, nucleic acid homology can be
determined through hybridisation studies. Nucleic acids
which hybridise under stringent conditions to the nucleic
20   acids of the invention are considered breast cancer
sequences. Under stringent conditions, hybridisation will
most preferably occur at 42°C in 750 mM NaCl, 75 mM
trisodium citrate, 2% SDS, 50% formamide, 1X Denhart's,
10% (w/v) dextran sulphate and 100 µg/ml denatured salmon
25   sperm DNA. Useful variations on these conditions will be
readily apparent to those skilled in the art. The washing
steps which follow hybridization most preferably occur at
65°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 1% SDS.
Additional variations on these conditions will be readily
30   apparent to those skilled in the art.

In a further aspect, the invention provides breast
cancer sequences as described in Table 1 and represented
by the SEQ ID Numbers: 1-11, or the nucleotide sequence of
a nucleic acid which hybridises thereto as described
35   above, and appropriate control elements of the breast
cancer sequences.

Preferably the control elements are those which

mediate expression in breast tissue, but may also mediate expression in other tissues including, but not restricted to, prostate, liver and ovary. The breast cancer nucleic acid sequences of the present invention can be engineered using methods accepted in the art so as to alter the sequences for a variety of purposes. These include, but are not limited to, modification of the cloning, processing, and/or expression of the gene product. PCR reassembly of gene fragments and the use of synthetic oligonucleotides allow the engineering of breast cancer sequences of the invention. For example, oligonucleotide-mediated site-directed mutagenesis can introduce mutations that create new restriction sites, alter glycosylation patterns and produce splice variants etc.

As a result of the degeneracy of the genetic code, a number of polynucleotide sequences encoding breast cancer proteins of the invention, some that may have minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention includes each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring breast cancer sequences, and all such variations are to be considered as being specifically disclosed.

The polynucleotides of this invention include RNA, cDNA, genomic DNA, synthetic forms, and mixed polymers, both sense and antisense strands, and may be chemically or biochemically modified, or may contain non-natural or derivatised nucleotide bases as will be appreciated by those skilled in the art. Such modifications include labels, methylation, intercalators, alkylators and modified linkages. In some instances it may be advantageous to produce nucleotide sequences encoding breast cancer sequences of the invention, or their

- 8 -

derivatives, possessing a substantially different codon usage than that of the naturally occurring gene. For example, codons may be selected to increase the rate of expression of the peptide in a particular prokaryotic or eukaryotic host corresponding with the frequency that particular codons are utilized by the host. Other reasons to alter the nucleotide sequence encoding breast cancer sequences of the invention, or their derivatives, without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

In some instances the breast cancer nucleic acid sequences of the invention are fragments of larger genes and may be used to identify and obtain corresponding full-length genes. Full-length sequences of the breast cancer genes can be obtained using the partial gene sequences, such as BNO8, BNO205 and BNO221 described in Table 1, by methods known per se to those skilled in the art. For example, "restriction-site PCR" may be used to retrieve unknown sequence adjacent to a portion of DNA whose sequence is known. In this technique universal primers are used to retrieve unknown sequence. Inverse PCR may also be used, in which primers based on the known sequence are designed to amplify adjacent unknown sequences. These upstream sequences may include promoters and regulatory elements. In addition, various other PCR-based techniques may be used, for example a kit available from Clontech (Palo Alto, California) allows for a walking PCR technique, the 5'RACE kit (Gibco-BRL) allows isolation of additional 5' gene sequence while additional 3' sequence can be obtained using practised techniques (for eg see Gecz et al., 1997).

The invention also encompasses production of breast cancer sequences of the invention entirely by synthetic chemistry. Synthetic sequences may be inserted into expression vectors and cell systems that contain the

necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. Numerous types of appropriate expression vectors and suitable regulatory elements are known in the art for a
5   variety of host cells. Regulatory elements may include regulatory sequences, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, 5' and 3' untranslated regions and specific translational start and stop signals (such as an ATG initiation codon
10   and Kozak consensus sequence). Regulatory elements will allow more efficient translation of sequences encoding breast cancer genes of the invention. In cases where the complete coding sequence including the initiation codon and upstream regulatory sequences are inserted into the
15   appropriate expression vector, additional control signals may not be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals as described above should be provided by the vector. Such signals may be of various
20   origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used (Scharf et al., 1994).

The present invention allows for the preparation of
25   purified breast cancer polypeptide or protein, from the polynucleotides of the present invention or variants thereof. In order to do this, host cells may be transfected with a nucleic acid molecule as described above. Typically said host cells are transfected with an
30   expression vector comprising a nucleic acid encoding a breast cancer protein according to the invention. Cells are cultured under the appropriate conditions to induce or cause expression of the breast cancer protein. The conditions appropriate for breast cancer protein
35   expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art.

A variety of expression vector/host systems may be
utilized to contain and express the breast cancer
sequences of the invention and are well known in the art.
These include, but are not limited to, microorganisms such

5   as bacteria transformed with plasmid or cosmid DNA
expression vectors; yeast transformed with yeast
expression vectors; insect cell systems infected with
viral expression vectors (e.g., baculovirus); or mouse or
other animal or human tissue cell systems. In a preferred

10   embodiment the breast cancer proteins of the invention are
expressed in mammalian cells using various expression
vectors including plasmid, cosmid and viral systems such
as adenoviral, retroviral or vaccinia virus expression
systems. The invention is not limited by the host cell

15   employed.

The polynucleotide sequences, or variants thereof, of
the present invention can be stably expressed in cell
lines to allow long term production of recombinant
proteins in mammalian systems.   These sequences can be

20   transformed into cell lines using expression vectors which
may contain viral origins of replication and/or endogenous
expression elements and a selectable marker gene on the
same or on a separate vector. The selectable marker
confers resistance to a selective agent, and its presence

25   allows growth and recovery of cells which successfully
express the introduced sequences. Resistant clones of
stably transformed cells may be propagated using tissue
culture techniques appropriate to the cell type.

The protein produced by a transformed cell may be

30   secreted or retained intracellularly depending on the
sequence and/or the vector used. As will be understood by
those of skill in the art, expression vectors containing
polynucleotides which encode a protein of the invention
may be designed to contain signal sequences which direct

35   secretion of the protein through a prokaryotic or
eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its

ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, glycosylation,
5   phosphorylation, and acylation. Post-translational cleavage of a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells having specific cellular machinery and characteristic mechanisms for post-
10   translational activities (e.g., CHO or HeLa cells), are available from the American Type Culture Collection (ATCC) and may be chosen to ensure the correct modification and processing of the foreign protein.

When large quantities of protein are needed such as
15   for antibody production, vectors which direct high levels of breast cancer gene expression may be used such as those containing the T5 or T7 inducible bacteriophage promoter. The present invention also includes the use of the expression systems described above in generating and
20   isolating fusion proteins which contain important functional domains of the protein. These fusion proteins are used for binding, structural and functional studies as well as for the generation of appropriate antibodies.

In order to express and purify the protein as a
25   fusion protein, the appropriate cDNA sequence is inserted into a vector which contains a nucleotide sequence encoding another peptide (for example, glutathionine succinyl transferase). The fusion protein is expressed and recovered from prokaryotic or eukaryotic cells. The fusion
30   protein can then be purified by affinity chromatography based upon the fusion vector sequence. The relevant protein can subsequently be obtained by enzymatic cleavage of the fusion protein.

In one embodiment, a fusion protein may be generated
35   by the fusion of a breast cancer polypeptide with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is

- 12 -

generally placed at the amino- or carboxy-terminus of the breast cancer polypeptide. The presence of such epitope-tagged forms of a breast cancer polypeptide can be detected using an antibody against the tag polypeptide.

5   Also, provision of the epitope tag enables the breast cancer polypeptide to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag.

Various tag polypeptides and their respective

10 antibodies are well known in the art. Examples include poly-histidine or poly-histidine-glycine tags and the c-myc tag and antibodies thereto.

Fragments of breast cancer polypeptide may also be produced by direct peptide synthesis using solid-phase

15 techniques. Automated synthesis may be achieved by using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of breast cancer polypeptide may be synthesized separately and then combined to produce the full-length molecule.

20   In a further aspect of the invention there is provided a method of preparing a polypeptide as described above, comprising the steps of:

(1) culturing the host cells under conditions effective for production of the polypeptide; and

25   (2) harvesting the polypeptide.

Substantially purified breast cancer proteins or fragments thereof can then be used in further biochemical analyses to establish secondary and tertiary structure for example by x-ray crystallography of the protein or by

30 nuclear magnetic resonance (NMR). Determination of structure allows for the rational design of pharmaceuticals to interact with the protein, alter protein charge configuration or charge interaction with other proteins, or to alter its function in the cell.

35   The breast cancer sequences of the present invention have been identified from a region of restricted LOH seen in breast cancer. In addition, these breast cancer genes

- 13 -

have been shown to be differentially expressed in breast
cancer samples compared with normal tissue controls. As
LOH is suggestive of the presence of a tumour suppressor
gene, those breast cancer genes of the invention that are
5   down-regulated in their expression in cancerous tissue, as
highlighted in Figures 2 and 3 and listed in SEQ ID
Numbers: 1-9 represent tumour suppressor genes in the
16q24.3 region. As many of these genes are expressed in a
wide variety of tissues and LOH of 16q has been found in
10   cancers of other tissue types, including prostate, liver,
ovary, primitive neuroectodermal and Wilms' tumours, they
may represent tumour suppressor genes involved in a range
of cancers. Such cancers may include, but are not limited
to adenocarcinoma, leukemia, lymphoma, melanoma, myeloma,
15   sarcoma, teratocarcinoma, and, in particular, cancers of
the breast, prostate, blood, germ cells, liver, ovary,
adrenal gland, cervix, heart, brain, lung, placenta,
skeletal muscle, synovial membrane, tonsil, lymph tissue,
kidney, colon, uterus, skin and testis. Other cancers may
20   include those of the head and neck, bladder, bone, bone
marrow, gall bladder, ganglia, gastrointestinal tract,
pancreas, parathyroid, penis, salivary glands, spleen,
stomach, thymus and thyroid gland.

In addition, the identification of BNO223 (SEQ ID
25   Numbers: 10 and 11) to be up-regulated in its expression
in breast cancer samples suggests a role as an oncogene.
This gene is also expressed in many tissue types and as
such may be a causative factor in other cancers such as
those listed above.

30   With the identification of the breast cancer
nucleotide and protein sequences of the invention, probes
and antibodies raised to the genes can be used in a
variety of hybridisation and immunological assays to
screen for and detect the presence of either a normal or
35   mutated gene or gene product.

In addition the_nucleotide and protein sequences of
the breast cancer genes provided in this invention enable

- 14 -

therapeutic methods for the treatment of cancers associated with one or more of these genes, enable screening of compounds for therapeutic intervention, and also enable methods for the diagnosis or prognosis of all cancers associated with the these genes. Examples of such cancers include, but are not limited to, those listed above.

In the treatment of cancers associated with down-regulated gene expression and/or activity, it is desirable to increase the expression and/or activity of the relevant gene. In the treatment of disorders associated with up-regulated gene expression and/or activity, it is desirable to decrease the expression and/or activity of the relevant gene.

Enhancing breast cancer gene or protein function

Enhancing, stimulating or re-activating the function of those breast cancer genes or proteins that are down-regulated in cancer can be achieved in a variety of ways as would be appreciated by those skilled in the art.

In a preferred embodiment a breast cancer gene of the invention is administered to a subject to treat or prevent a cancer associated with decreased activity and/or expression of the gene.

In a further aspect, there is provided the use of a nucleic acid molecule of the invention, as described above, in the manufacture of a medicament for the treatment of a cancer associated with decreased activity and/or expression of the corresponding gene.

Typically, a vector capable of expressing a breast cancer gene of the invention, or fragment or derivative thereof, may be administered to a subject to treat or prevent a cancer associated with decreased activity and/or expression of the gene, including but not limited to, those described above.

Transducing retroviral vectors are often used for somatic cell gene therapy because of their high efficiency

of infection and stable integration and expression. The
full-length breast cancer gene, or portions thereof, can
be cloned into a retroviral vector and expression can be
driven from its endogenous promoter or from the retroviral
5   long terminal repeat or from a promoter specific for the
target cell type of interest. Other viral vectors can be
used and include, as is known in the art, adenoviruses,
adeno-associated virus, vaccinia virus, papovaviruses,
lentiviruses and retroviruses of avian, murine and human
10   origin.

Gene therapy would be carried out according to
established methods (Friedman, 1991; Culver, 1996). A
vector containing a copy of a breast cancer gene linked to
expression control elements and capable of replicating
15   inside the cells is prepared. Alternatively the vector may
be replication deficient and may require helper cells or
helper virus for replication and virus production and use
in gene therapy.

Gene transfer using non-viral methods of infection
20   can also be used. These methods include direct injection
of DNA, uptake of naked DNA in the presence of calcium
phosphate, electroporation, protoplast fusion or liposome
delivery. Gene transfer can also be achieved by delivery
as a part of a human artificial chromosome or receptor-
25   mediated gene transfer. This involves linking the DNA to
a targeting molecule that will bind to specific cell-
surface receptors to induce endocytosis and transfer of
the DNA into mammalian cells. One such technique uses
poly-L-lysine to link asialoglycoprotein to DNA.   An
30   adenovirus is also added to the complex to disrupt the
lysosomes and thus allow the DNA to avoid degradation and
move   to   the   nucleus. Infusion   of   these   particles
intravenously   has   resulted   in   gene   transfer   into
hepatocytes.

35   In affected subjects that express a mutated form of a
breast cancer gene of the invention, it may be possible to
prevent the cancer by introducing into the affected cells

a wild-type copy of the gene such that it recombines with the mutant gene. This requires a double recombination event for the correction of the gene mutation. Vectors for the introduction of genes in these ways are known in the art, and any suitable vector may be used. Alternatively, introducing another copy of the gene bearing a second mutation in that gene may be employed so as to negate the original gene mutation and block any negative effect.

In a still further aspect the invention provides a method for the treatment of a cancer associated with decreased activity and/or expression of a breast cancer gene of the invention, comprising administering a polypeptide as described above, or an agonist thereof, to a subject in need of such treatment.

In another aspect the invention provides the use of a polypeptide as described above, or an agonist thereof, in the manufacture of a medicament for the treatment of a cancer associated with decreased activity and/or expression of a breast cancer gene.

In affected subjects that have decreased expression of a breast cancer gene, a mechanism of down-regulation may be abnormal methylation of a CpG island if present in the 5' end of the gene. Therefore, in an alternative approach to therapy, administration of agents that remove breast cancer gene promoter methylation will reactivate its expression which may suppress the associated cancer phenotype.

Inhibiting breast cancer gene or protein function

Inhibiting the function of those breast cancer genes or proteins of the invention that are up-regulated in cancer can be achieved in a variety of ways as would be appreciated by those skilled in the art.

In one aspect of the invention there is provided a method of treating a cancer associated with increased activity and/or expression of a breast cancer gene, comprising administering an antagonist of the gene to a

- 17 -

subject in need of such treatment.

In still another aspect of the invention there is provided the use of an antagonist of a breast cancer gene in the manufacture of a medicament for the treatment of a
5   cancer associated with increased activity and/or expression of the gene.

In one aspect of the invention an isolated DNA molecule, which is the complement of any one of the DNA molecules described above and which encodes an RNA
10   molecule that hybridises with the mRNA encoded by a breast cancer gene of the invention, may be administered to a subject in need of such treatment.

In a still further aspect of the invention there is provided the use of an isolated DNA molecule which is the
15   complement of a nucleic acid molecule of the invention and which encodes an RNA molecule that hybridises with the mRNA encoded by a breast cancer gene, in the manufacture of a medicament for the treatment of a disorder associated with increased activity and/or expression of the gene.

20   Typically, a vector expressing the complement of a polynucleotide encoding a breast cancer gene of the invention may be administered to a subject to treat or prevent a disorder associated with increased activity and/or expression of the gene including, but not limited
25   to, those described above. Antisense strategies may use a variety of approaches including the use of antisense oligonucleotides, ribozymes, DNAzymes, injection of antisense RNA and transfection of antisense RNA expression vectors. Many methods for introducing vectors into cells
30   or tissues are available and equally suitable for use *in vivo*, *in vitro*, and *ex vivo*. For *ex vivo* therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by
35   liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (For example, see Goldman *et al.*, 1997).

- 18 -

According to still another aspect of the invention, there is provided a method of treating a cancer associated with increased activity and/or expression of a breast cancer gene of the invention comprising administering an antagonist of the gene to a subject in need of such treatment.

In still another aspect of the invention there is provided the use of an antagonist of a breast cancer gene of the invention in the manufacture of a medicament for the treatment of a cancer associated with increased activity and/or expression of the gene.

Such disorders may include, but are not limited to, those discussed above. In one aspect purified protein according to the invention may be used to produce antibodies which specifically bind the breast cancer protein. These antibodies may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues that express the protein. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric and single chain antibodies as would be understood by the person skilled in the art.

For the production of antibodies, various hosts including rabbits, rats, goats, mice, humans, and others may be immunized by injection with a protein of the invention or with any fragment or oligopeptide thereof, which has immunogenic properties. Various adjuvants may be used to increase immunological response and include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface-active substances such as lysolecithin. Adjuvants used in humans include BCG (bacilli Calmette-Guerin) and Corynebacterium parvum.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to the breast cancer proteins of the invention have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also

preferable that these oligopeptides, peptides, or
fragments are identical to a portion of the amino acid
sequence of the natural protein and contain the entire
amino acid sequence of a small, naturally occurring
5   molecule. Short stretches of amino acids from these
proteins may be fused with those of another protein, such
as KLH, and antibodies to the chimeric molecule may be
produced.

Monoclonal antibodies to breast cancer proteins of
10 the invention may be prepared using any technique which
provides for the production of antibody molecules by
continuous cell lines in culture. These include, but are
not limited to, the hybridoma technique, the human B-cell
hybridoma technique, and the EBV-hybridoma technique. (For
15 example, see Kohler *et al.*, 1975; Kozbor *et al.*, 1985;
Cote *et al.*, 1983; Cole *et al.*, 1984).

Antibodies may also be produced by inducing *in vivo*
production in the lymphocyte population or by screening
immunoglobulin libraries or panels of highly specific
20 binding reagents as disclosed in the literature. (For
example, see Orlandi *et al.*, 1989; Winter *et al.*, 1991).

Antibody fragments which contain specific binding
sites for the breast cancer proteins may also be
generated. For example, such fragments include, F(ab')2
25 fragments produced by pepsin digestion of the antibody
molecule and Fab fragments generated by reducing the
disulfide bridges of the F(ab')2 fragments. Alternatively,
Fab expression libraries may be constructed to allow rapid
and easy identification of monoclonal Fab fragments with
30 the desired specificity. (For example, see Huse *et al.*,
1989).

Various immunoassays may be used for screening to
identify antibodies having the desired specificity.
Numerous protocols for competitive binding or
35 immunoradiometric assays using either polyclonal or
monoclonal antibodies with established specificities are
well known in the art. Such immunoassays typically involve

the measurement of complex formation between a protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes is preferred, but a competitive binding assay may also be employed.

Drug screening

According to still another aspect of the invention, the breast cancer nucleic acids and proteins of the invention, and cells expressing these, are useful for screening of candidate pharmaceutical agents or compounds in a variety of techniques for the treatment of cancers associated with their dysfunction.

Candidate pharmaceutical agents or compounds encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having molecular weight of more than 100 and less than about 2,500 daltons. Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids and steroids. Particularly preferred are peptides.

Agent screening techniques include, but are not limited to, utilising eukaryotic or prokaryotic host cells that are stably transformed with recombinant molecules expressing a particular breast cancer polypeptide of the invention, or fragment thereof, preferably in competitive binding assays. Binding assays will measure for the formation of complexes between the breast cancer polypeptide, or fragments thereof, and the agent being tested, or will measure the degree to which an agent being tested will interfere with the formation of a complex between the breast cancer polypeptide, or fragment thereof, and a known ligand.

Another technique for drug screening provides high-throughput screening for compounds having suitable binding affinity to a breast cancer polypeptide (see PCT published application WO84/03564). In this stated technique, large numbers of small peptide test compounds can be synthesised

on a solid substrate and can be assayed through breast cancer polypeptide binding and washing. Bound breast cancer polypeptide is then detected by methods well known in the art. In a variation of this technique, purified

5   polypeptides can be coated directly onto plates to identify interacting test compounds.

An additional method for drug screening involves the use of host eukaryotic cell lines which carry mutations in a particular breast cancer gene. The host cell lines are

10   also defective at the polypeptide level. Other cell lines may be used where the gene expression of the breast cancer gene can be switched off or up-regulated. The host cell lines or cells are grown in the presence of various drug compounds and the rate of growth of the host cells is

15   measured to determine if the compound is capable of regulating the growth of defective cells.

Breast cancer polypeptide may also be used for screening compounds developed as a result of combinatorial library technology. This provides a way to test a large

20   number of different substances for their ability to modulate activity of a polypeptide. The use of peptide libraries is preferred (see patent WO97/02048) with such libraries and their use known in the art.

A substance identified as a modulator of polypeptide

25   function may be peptide or non-peptide in nature. Non-peptide "small molecules" are often preferred for many *in vivo* pharmaceutical applications. In addition, a mimic or mimetic of the substance may be designed for pharmaceutical use. The design of mimetics based on a

30   known pharmaceutically active compound ("lead" compound) is a common approach to the development of novel pharmaceuticals. This is often desirable where the original active compound is difficult or expensive to synthesise or where it provides an unsuitable method of

35   administration. In the design of a mimetic, particular parts of the original active compound that are important in determining the target property are identified. These

parts or residues constituting the active region of the compound are known as its pharmacophore. Once found, the pharmacophore structure is modelled according to its physical properties using data from a range of sources including x-ray diffraction data and NMR. A template molecule is then selected onto which chemical groups which mimic the pharmacophore can be added. The selection can be made such that the mimetic is easy to synthesise, is likely to be pharmacologically acceptable, does not degrade *in vivo* and retains the biological activity of the lead compound. Further optimisation or modification can be carried out to select one or more final mimetics useful for *in vivo* or clinical testing.

It is also possible to isolate a target-specific antibody and then solve its crystal structure. In principle, this approach yields a pharmacophore upon which subsequent drug design can be based as described above. It may be possible to avoid protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analogue of the original binding site. The anti-id could then be used to isolate peptides from chemically or biologically produced peptide banks.

In further embodiments, any of the genes, proteins, antagonists, antibodies, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents may be made by those skilled in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, therapeutic efficacy with lower dosages of each agent may be possible, thus reducing the potential for adverse side effects.

In a further aspect a pharmaceutical composition and a pharmaceutically acceptable carrier may be administered. The pharmaceutical composition may comprise any one or more of a polypeptide as described above, typically a
5   substantially purified breast cancer polypeptide, an antibody to a breast cancer polypeptide, a vector capable of expressing a breast cancer polypeptide, a compound which increases or decreases expression of a breast cancer gene, a candidate drug that restores wild-type activity to
10   a breast cancer gene or an antagonist of a breast cancer gene.

The pharmaceutical composition may be administered to a subject to treat or prevent a cancer associated with decreased activity and/or expression of a breast cancer
15   gene including, but not limited to, those provided above. Pharmaceutical compositions in accordance with the present invention are prepared by mixing a polypeptide of the invention, or active fragments or variants thereof, having the desired degree of purity, with acceptable carriers,
20   excipients, or stabilizers which are well known. Acceptable carriers, excipients or stabilizers are nontoxic at the dosages and concentrations employed, and include buffers such as phosphate, citrate, and other organic acids; antioxidants including absorbic acid; low
25   molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine;
30   monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitrol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as Tween, Pluronics or
35   polyethylene glycol (PEG).

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including,

- 24 -

for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

Diagnostic and prognostic applications

Polynucleotide sequences encoding the breast cancer genes of the invention may be used for the diagnosis or prognosis of cancers associated with their dysfunction, or a predisposition to such cancers. Examples of such cancers include, but are not limited to, adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the breast, prostate, blood, germ cells, liver, ovary, adrenal gland, cervix, heart, brain, lung, placenta, skeletal muscle, synovial membrane, tonsil, lymph tissue, kidney, colon, uterus, skin and testis. Other cancers may include those of the head and neck, bladder, bone, bone marrow, gall bladder, ganglia, gastrointestinal tract, pancreas, parathyroid, penis, salivary glands, spleen, stomach, thymus and thyroid gland.

Diagnosis or prognosis may be used to determine the severity, type or stage of the disease state in order to initiate an appropriate therapeutic intervention.

In another embodiment of the invention, the polynucleotides that may be used for diagnostic or prognostic purposes include oligonucleotide sequences, genomic DNA and complementary RNA and DNA molecules. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which mutations or abnormal expression of the relevant breast cancer gene may be correlated with disease. Genomic DNA used for the diagnosis or prognosis may be obtained from body cells, such as those present in the blood, tissue biopsy, surgical specimen, or autopsy material. The DNA may be isolated and used directly for detection of a specific sequence or may be amplified by the polymerase chain reaction (PCR) prior to analysis. Similarly, RNA or cDNA may also be used, with or without PCR amplification. To

detect a specific nucleic acid sequence, direct nucleotide sequencing, reverse transcriptase PCR (RT-PCR), hybridization using specific oligonucleotides, restriction enzyme digest and mapping, PCR mapping, RNAse protection,

5   and various other methods may be employed. Oligonucleotides specific to particular sequences can be chemically synthesized and labelled radioactively or non-radioactively and hybridised to individual samples immobilized on membranes or other solid-supports or in

10   solution. The presence, absence or excess expression of a particular breast cancer gene may then be visualized using methods such as autoradiography, fluorometry, or colorimetry.

     In a particular aspect, the nucleotide sequences

15   encoding a breast cancer gene of the invention may be useful in assays that detect the presence of associated disorders, particularly those mentioned previously. The nucleotide sequences encoding the relevant breast cancer gene may be labelled by standard methods and added to a

20   fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is

25   significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding the breast cancer gene in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a

30   particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

     In order to provide a basis for the diagnosis or prognosis of a disorder associated with a mutation in a

35   particular breast cancer gene of the invention, the nucleotide sequence of the relevant gene can be compared between normal tissue and diseased tissue in order to

establish whether the patient expresses a mutant gene.

In order to provide a basis for the diagnosis or prognosis of a disorder associated with abnormal expression of a particular breast cancer gene of the

5   invention, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding the relevant breast cancer gene, under conditions

10 suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Another method to

15 identify a normal or standard profile for expression of a particular breast cancer gene is through quantitative RT-PCR studies. RNA isolated from body cells of a normal individual, particularly RNA isolated from tumour cells, is reverse transcribed and real-time PCR using

20 oligonucleotides specific for the relevant breast cancer gene is conducted to establish a normal level of expression of the gene.

Standard values obtained in both these examples may be compared with values obtained from samples from

25 patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays or

30 quantitative RT-PCR studies may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over

35 a period ranging from several days to months.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences,

including genomic sequences, encoding a particular breast cancer gene, or closely related molecules, may be used to identify nucleic acid sequences which encode the gene. The specificity of the probe, whether it is made from a highly

5   specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding the breast cancer gene,

10   allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the breast cancer encoding sequences. The hybridization probes of the subject

15   invention may be DNA or RNA and may be derived from the sequence of SEQ ID Numbers: 1-11 or from genomic sequences including promoters, enhancers, and introns of the genes.

Means for producing specific hybridization probes for DNAs encoding the breast cancer genes of the invention

20   include the cloning of polynucleotide sequences encoding these genes or their derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, and are commercially available. Hybridization probes may be labelled by radionuclides such as ^{32}P or ^{35}S, or by

25   enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, or other methods known in the art.

According to a further aspect of the invention there is provided the use of a polypeptide as described above in

30   the diagnosis or prognosis of a cancer associated with a breast cancer gene of the invention, or a predisposition to such cancers.

When a diagnostic or prognostic assay is to be based upon a breast cancer protein, a variety of approaches are

35   possible. For example, diagnosis or prognosis can be achieved by monitoring differences in the electrophoretic mobility of normal and mutant proteins. Such an approach

will be particularly useful in identifying mutants in
which charge substitutions are present, or in which
insertions, deletions or substitutions have resulted in a
significant change in the electrophoretic migration of the
5   resultant protein. Alternatively, diagnosis may be based
upon differences in the proteolytic cleavage patterns of
normal and mutant proteins, differences in molar ratios of
the various amino acid residues, or by functional assays
demonstrating altered function of the gene products.

10   In another aspect, antibodies that specifically bind
a breast cancer gene of the invention may be used for the
diagnosis or prognosis of cancers characterized by
abnormal expression of the gene, or in assays to monitor
patients being treated with the gene or agonists,
15   antagonists, or inhibitors of the gene. Antibodies useful
for diagnostic purposes may be prepared in the same manner
as described above for therapeutics. Diagnostic or
prognostic assays include methods that utilize the
antibody and a label to detect a breast cancer gene of the
20   invention in human body fluids or in extracts of cells or
tissues. The antibodies may be used with or without
modification, and may be labelled by covalent or non-
covalent attachment of a reporter molecule.

A variety of protocols for measuring a breast cancer
25   gene of the invention, including ELISAs, RIAs, and FACS,
are known in the art and provide a basis for diagnosing
altered or abnormal levels of their expression. Normal or
standard values for their expression are established by
combining body fluids or cell extracts taken from normal
30   mammalian subjects, preferably human, with antibody to the
breast cancer protein under conditions suitable for
complex formation. The amount of standard complex
formation may be quantitated by various methods,
preferably by photometric means. Quantities of any of the
35   breast cancer genes expressed in subject, control, and
disease samples from biopsied tissues are compared with
the standard values. Deviation between standard and

subject values establishes the parameters for diagnosing disease.

Once an individual has been diagnosed with a cancer, effective treatments can be initiated. These may include
5   administering a selective agonist to the relevant mutant breast cancer gene so as to restore its function to a normal level or introduction of the wild-type gene, particularly through gene therapy approaches as described above. Typically, a vector capable of expressing the
10   appropriate full-length breast cancer gene or a fragment or derivative thereof may be administered. In an alternative approach to therapy, a substantially purified breast cancer polypeptide and a pharmaceutically acceptable carrier may be administered, as described
15   above, or drugs which can replace the function of or mimic the action of the relevant breast cancer gene may be administered.

In the treatment of cancers associated with increased breast cancer gene expression and/or activity, the
20   affected individual may be treated with a selective antagonist such as an antibody to the relevant protein or an antisense (complement) probe to the corresponding gene as described above, or through the use of drugs which may block the action of the relevant breast cancer gene.
25

Microarray

In further embodiments, complete cDNAs, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used
30   as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic
35   basis of a disorder, to diagnose or prognose a disorder, and to develop and monitor the activities of therapeutic agents. Microarrays may be prepared, used, and analyzed

using methods known in the art. (For example, see Schena
et al., 1996; Heller et al., 1997).

Transformed hosts

5   The present invention also provides for the
production of genetically modified (knock-out, knock-in
and transgenic), non-human animal models transformed with
the DNA molecules of the invention. These animals are
useful for the study of breast cancer gene function, to
10   study the mechanisms of cancer as related to the breast
cancer genes, for the screening of candidate
pharmaceutical compounds, for the creation of explanted
mammalian cell cultures which express the protein or
mutant protein and for the evaluation of potential
15   therapeutic interventions.

        One of the breast cancer genes of the invention may
have been inactivated by knock-out deletion, and knock-out
genetically modified non-human animals are therefore
provided.

20   Animal species which are suitable for use in the
animal models of the present invention include, but are
not limited to, rats, mice, hamsters, guinea pigs,
rabbits, dogs, cats, goats, sheep, pigs, and non-human
primates such as monkeys and chimpanzees. For initial
25   studies, genetically modified mice and rats are highly
desirable due to their relative ease of maintenance and
shorter life spans. For certain studies, transgenic yeast
or invertebrates may be suitable and preferred because
they allow for rapid screening and provide for much easier
30   handling. For longer term studies, non-human primates may
be desired due to their similarity with humans.

        To create an animal model for a mutated breast cancer
gene of the invention several methods can be employed.
These include generation of a specific mutation in a
35   homologous animal gene, insertion of a wild type human
gene and/or a humanized animal gene by homologous
recombination, insertion of a mutant (single or multiple)

human gene as genomic or minigene cDNA constructs using
wild type or mutant or artificial promoter elements or
insertion of artificially modified fragments of the
endogenous gene by homologous recombination. The
5   modifications include insertion of mutant stop codons, the
deletion of DNA sequences, or the inclusion of
recombination elements (lox p sites) recognized by enzymes
such as Cre recombinase.

To create a transgenic mouse, which is preferred, a
10 mutant version of a particular breast cancer gene of the
invention can be inserted into a mouse germ line using
standard techniques of oocyte microinjection or
transfection or microinjection into embryonic stem cells.
Alternatively, if it is desired to inactivate or replace
15 the endogenous breast cancer gene, homologous
recombination using embryonic stem cells may be applied.

For oocyte injection, one or more copies of the
mutant or wild type breast cancer gene can be inserted
into the pronucleus of a just-fertilized mouse oocyte.
20 This oocyte is then reimplanted into a pseudo-pregnant
foster mother. The liveborn mice can then be screened for
integrants using analysis of tail DNA for the presence of
human breast cancer gene sequences. The transgene can be
either a complete genomic sequence injected as a YAC, BAC,
25 PAC or other chromosome DNA fragment, a cDNA with either
the natural promoter or a heterologous promoter, or a
minigene containing all of the coding region and other
elements found to be necessary for optimum expression.

According to still another aspect of the invention
30 there is provided the use of genetically modified non-
human animals as described above for the screening of
candidate pharmaceutical compounds.

It will be clearly understood that, although a number
of prior art publications are referred to herein, this
35 reference does not constitute an admission that any of
these documents forms part of the common general knowledge
in the art, in Australia or in any other country.

- 32 -

Throughout this specification and the claims, the words "comprise", "comprises" and "comprising" are used in a non-exclusive sense, except where the context requires otherwise.

Brief Description of the Drawings

Figure 1. Schematic representation of tumours with interstitial and terminal allelic loss on chromosome arm 16q in the two series of tumour samples. Polymorphic markers are listed according to their order on 16q from centromere to telomere and the markers used for each series are indicated by X. Tumour identification numbers are shown at the top of each column. At the right of the figure, the three smallest regions of loss of heterozygosity are indicated.

Figure 2. Relative fold expression variability index (RFVI) for genes mapping to the 16q24.3 LOH region. Genes exhibiting an RFVI greater than 50 (Represented by hatched bars) were considered to be significantly differentially expressed in breast cancer cell line samples compared to normal breast tissue. The control tumour suppressor expression profiles for SYK and INK4A/ARF are also shown (Spotted bars).

Figure 3. Fold change of expression for genes mapping to the 16q24.3 LOH region in breast cancer cell line mRNA relative to normal breast tissue expression. For each gene, the percentage of cell lines exhibiting various fold differences in expression is indicated.

Modes for performing the invention

EXAMPLE 1: Collection of breast cancer patient material

Two series of breast cancer patients were analysed for this study. Histopathological classification of each tumour specimen was carried out by our collaborators according to World Health Organisation criteria (WHO, 1981). Patients were graded histopathologically according to the modified Bloom and Richardson method (Elston and

Ellis, 1990) and patient material was obtained upon approval of local Medical Ethics Committees. Tumour tissue DNA and peripheral blood DNA from the same individual was isolated as previously described (Devilee et al., 1991)
5 using standard laboratory protocols.

Series 1 consisted of 189 patients operated on between 1986 and 1993 in three Dutch hospitals, a Dutch University and two peripheral centres. Tumour tissue was snap frozen within a few hours of resection. For DNA
10 isolation, a tissue block was selected only if it contained at least 50% of tumour cells following examination of haematoxilin and eosin stained tissue sections by a pathologist. Tissue blocks that contained fewer than 50% of tumour cells were omitted from further
15 analysis.

Series 2 consisted of 123 patients operated on between 1987 and 1997 at the Flinders Medical Centre in Adelaide, Australia. Of these, 87 were collected as fresh specimens within a few hours of surgical resection,
20 confirmed as malignant tissue by pathological analysis, snap frozen in liquid nitrogen, and stored at -70°C. The remaining 36 tumour tissue samples were obtained from archival paraffin embedded tumour blocks. Prior to DNA isolation, tumour cells were microdissected from tissue
25 sections mounted on glass slides so as to yield at least 80% tumour cells. In some instances, no peripheral blood was available such that pathologically identified paraffin embedded non-malignant lymph node tissue was used instead.

30 EXAMPLE 2: LOH analysis of chromosome 16q markers in breast cancer samples.

In order to identify the location of genes associated with breast cancer, LOH analysis of tumour samples was conducted. A total of 45 genetic markers mapping to
35 chromosome 16 were used for the LOH analysis of the breast tumour and matched normal DNA samples collected for this study. Figure 1 indicates for which tumour series they

- 34 -

were used and their cytogenetic location. Details
regarding all markers can be obtained from the Genome
Database (GDB) at http://www.gdb.org. The physical order
of markers with respect to each other was determined from
5   a combination of information in GDB, by mapping on a
chromosome 16 somatic cell hybrid map (Callen et al.,
1995) and by genomic sequence information.

Four alternative methods were used for the LOH
analysis:

10   1)   For RFLP and VNTR markers, Southern blotting was
used to test for allelic imbalance. These markers were
used on only a subset of samples. Methods used were as
previously described (Devilee et al., 1991).

2)   Microsatellite markers were amplified from
15   tumour and normal DNA using the polymerase chain reaction
(PCR) incorporating standard methodologies (Weber and May,
1989; Sambrook et al., 1989). A typical reaction consisted
of 12 µl and contained 100 ng of template, 5 pmol of both
primers, 0.2 mM of each dNTP, 1 µCurie [α-^{32}P]dCTP, 1.5 mM
20   MgCl$_2$, 1.2 µl Supertaq buffer and 0.06 units of Supertaq
(HT biotechnologies). A Phosphor Imager type 445 SI
(Molecular Dynamics, Sunnyvale, CA) was used to quantify
ambiguous results. In these cases, the Allelic Imbalance
Factor (AIF) was determined as the quotient of the peak
25   height ratios from the normal and tumour DNA pair. The
threshold for allelic imbalance was defined as a 40%
reduction of one allele, agreeing with an AIF of $\geq$1.7 or
$\leq$0.59. This threshold is in accordance with the selection
of tumour tissue blocks containing at least 50% tumour
30   cells with a 10% error-range. The threshold for retention
has been previously determined to range from 0.76 to 1.3
(Devilee et al., 1994). This leaves a range of AIFs (0.58
- 0.75 and 1.31 - 1.69) for which no definite decision has
been made. This "grey area" is indicated by grey boxes in
35   Figure 1 and tumours with only "grey area" values were
discarded completely from the analysis.

3)   The third method for determining allelic

- 35 -

imbalance was similar to the second method above, however radioactively labelled dCTP was omitted. Instead, PCR of polymorphic microsatellite markers was done with one of the PCR primers labelled fluorescently with FAM, TET or
5   HEX. Analysis of PCR products generated was on an ABI 377 automatic sequencer (PE Biosystems) using 6% polyacrylamide gels containing 8M urea. Peak height values and peak sizes were analysed with the GeneScan programme (PE Biosystems). The same thresholds for allelic
10 imbalance, retention and grey areas were used as for the radioactive analysis.

        4)   An alternative fluorescent based system was also used. In this instance PCR primers were labelled with fluorescein or hexachlorofluorescein. PCR reaction volumes
15 were 20 µl and included 100 ng of template, 100 ng of each primer, 0.2 mM of each dNTP, 1-2 mM MgCl$_2$, 1X AmpliTaq Gold buffer and 0.8 units AmpliTaq Gold enzyme (Perkin Elmer). Cycling conditions were 10 cycles of 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 1 minute, followed by 25
20 cycles of 94°C 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, with a final extension of 72°C for 10 minutes. PCR amplimers were analysed on an ABI 373 automated sequencer (PE Biosystems) using the GeneScan programme (PE Biosystems). The threshold range of AIF for allele
25 retention was defined as 0.61 - 1.69, allelic loss as ≤0.5 or ≥2.0, or the "grey area" as 051 - 0.6 or 1.7 - 1.99.

        The first three methods were applied to the first tumour series while the last method was adopted for the second series of tumour samples. For statistical analysis,
30 a comparison of allelic imbalance data for validation of the different detection methods and of the different tumour series was done using the Chi-square test.

        The identification of the smallest region of overlap (SRO) involved in LOH is instrumental for narrowing down
35 the location of the gene targeted by LOH. Figure 1 shows the LOH results for tumour samples, which displayed small regions of loss (ie interstitial and telomeric LOH) and

does not include samples that showed complex LOH
(alternating loss and retention of markers). When
comparing the two sample sets at least three consistent
regions emerge with two being at the telomere in band

5   16q24.3 and one at 16q22.1. The region at 16q22.1 is
defined by the markers D16S398 and D16S301 and is based on
the interstitial LOH events seen in three tumours from
series 1 (239/335/478) and one tumour from series 2 (237).
At the telomere (16q24.2 - 16q24.3), the first region is

10   defined by the markers D16S498 and D16S3407 and is based
on four tumours from series 2 (443/75/631/408) while the
second region (16q24.3) extends from D16S3407 to the
telomere and is based on one tumour from series 1 (559)
and three from series 2 (97/240/466). LOH limited to the

15   telomere but involving both of the regions identified at
this site could be found in an additional 17 tumour
samples.

Other studies have shown that the long arm of
chromosome 16 is also a target for LOH in prostate, lung,

20   hepatocellular, ovarian, rhabdomyosarcoma and Wilms'
tumours. Detailed analysis of prostate carcinomas has
revealed an overlap in the smallest regions of LOH seen in
this cancer to that seen with breast cancer which suggests
that 16q harbours a gene implicated in many tumour types.

25

EXAMPLE 3: Construction of a physical map of 16q24.3

To identify novel candidate breast cancer genes
mapping to the smallest regions of overlap at 16q24.3, a
clone based physical map contig covering this region was

30   needed. At the start of this phase of the project the most
commonly used and readily accessible cloned genomic DNA
fragments were contained in lambda, cosmid or YAC vectors.
During the construction of whole chromosome 16 physical
maps, clones from a number of YAC libraries were

35   incorporated into the map (Doggett et al., 1995). These
included clones from a flow-sorted chromosome 16-specific
YAC library (McCormick et al., 1993), from the CEPH Mark I

and MegaYAC libraries and from a half-telomere YAC library
(Riethman *et al.*, 1989). Detailed STS and Southern
analysis of YAC clones mapping at 16q24.3 established that
very few were localised between the CY2/CY3 somatic cell
5   hybrid breakpoint and the long arm telomere. However,
those that were located in this region gave inconsistent
mapping results and were suspected to be rearranged or
deleted. Coupled with the fact that YAC clones make poor
sequencing substrates, and the difficulty in isolating the
10 cloned human DNA, a physical map based on cosmid clones
was the initial preferred option.

A flow-sorted chromosome 16 specific cosmid library
had previously been constructed (Longmire *et al.*, 1993),
with individual cosmid clones gridded in high-density
15 arrays onto nylon membranes. These filters collectively
contained ~15,000 clones representing an approximately 5.5
fold coverage of chromosome 16. Individual cosmids mapping
to the critical regions at 16q24.3 were identified by the
hybridisation of these membranes with markers identified
20 by this and previous studies to map to the region. The
strategy to align overlapping cosmid clones was based on
their STS content and restriction endonuclease digestion
pattern. Those clones extending furthest within each
initial contig were then used to walk along the chromosome
25 by the hybridisation of the ends of these cosmids back to
the high-density cosmid grids. This process continued
until all initial contigs were linked and therefore the
region defining the location of the breast cancer tumour
suppressor genes would be contained within the map.
30 Individual cosmid clones representing a minimum tiling
path in the contig were then used for the identification
of transcribed sequences by exon trapping, and for genomic
sequencing.

Chromosome 16 was sorted from the mouse/human somatic
35 cell hybrid CY18, which contains this chromosome as the
only human DNA, and *Sau*3A partially digested CY18 DNA was
ligated into the *Bam*HI cloning site of the cosmid sCOS-1

vector. All grids were hybridised and washed using methods described in Longmire *et al.* (1993). Briefly, the 10 filters were pre-hybridised in 2 large bottles for at least 2 hours in 20 ml of a solution containing 6X SSC; 10

5 mM EDTA (pH8.0); 10X Denhardt's; 1% SDS and 100 µg/ml denatured fragmented salmon sperm DNA at 65°C. Overnight hybridisations with [α-^{32}P]dCTP labelled probes were performed in 20 ml of fresh hybridisation solution at 65°C. Filters were washed sequentially in solutions of 2X SSC;

10 0.1% SDS (rinse at room temperature), 2X SSC; 0.1% SDS (room temperature for 15 minutes), 0.1X SSC; 0.1% SDS (room temperature for 15 minutes), and 0.1X SSC; 0.1% SDS (twice for 30 minutes at 50°C if needed). Membranes were exposed at -70°C for between 1 to 7 days.

15   Initial markers used for cosmid grid screening were those known to be located below the somatic cell hybrid breakpoints CY2/CY3 and the long arm telomere (Callen *et al.*, 1995). These included three genes, *CMAR*, *DPEP1*, and *MC1R*; the microsatellite marker D16S303; an end fragment

20 from the cosmid 317E5, which contains the *BBC1* gene; and four cDNA clones, yc81e09, yh09a04, D16S532E, and ScDNA-C113. The IMAGE consortium cDNA clone, yc81e09, was obtained through screening an arrayed normalised infant brain oligo-dT primed cDNA library (Soares *et al.*, 1994),

25 with the insert from cDNA clone ScDNA-A55. Both the ScDNA-A55 and ScDNA-C113 clones were originally isolated from a hexamer primed heteronuclear cDNA library constructed from the mouse/human somatic cell hybrid CY18 (Whitmore *et al.*, 1994). The IMAGE cDNA clone yh09a04 was identified from

30 direct cDNA selection of the cosmid 37B2 which was previously shown to map between the CY18A(D2) breakpoint and the 16q telomere. The EST, D16S532E, was also mapped to the same region. Subsequent to these initial screenings, restriction fragments representing the ends of

35 cosmids were used to identify additional overlapping clones.

- 39 -

Contig assembly was based on methods previously described (Whitmore et al., 1998). Later during the physical map construction, genomic libraries cloned into BAC or PAC vectors (Genome Systems or Rosewell Park Cancer

5   Institute) became available. These libraries were screened to aid in chromosome walking or when gaps that could not be bridged by using the cosmid filters were encountered. All BAC and PAC filters were hybridised and washed according to manufacturers recommendations. Initially,

10   membranes were individually pre-hybridised in large glass bottles for at least 2 hours in 20 ml of 6X SSC; 0.5% SDS; 5X Denhardt's; 100 µg/ml denatured salmon sperm DNA at 65°C. Overnight hybridisations with [α-^{32}P]dCTP labelled probes were performed at 65°C in 20 ml of a solution

15   containing 6X SSC; 0.5% SDS; 100 µg/ml denatured salmon sperm DNA. Filters were washed sequentially in solutions of 2X SSC; 0.5% SDS (room temperature 5 minutes), 2X SSC; 0.1% SDS (room temperature 15 minutes) and 0.1X SSC; 0.5% SDS (37°C 1 hour if needed). PAC or BAC clones identified

20   were aligned to the existing contig based on their restriction enzyme pattern or formed unique contigs which were extended by additional filter screens.

A high-density physical map consisting of cosmid, BAC and PAC clones has been established, which extends

25   approximately 3 Mb from the telomere of the long arm of chromosome 16. This contig extends beyond the CY2/CY3 somatic cell hybrid breakpoint and includes the 2 regions of minimal LOH identified at the 16q24.3 region in breast cancer samples. To date, a single gap of unknown size

30   exists in the contig and will be closed by additional contig extension experiments. The depth of coverage has allowed the identification of a minimal tiling path of clones which were subsequently used as templates for gene identification methods such as exon trapping and genomic

35   DNA sequencing.

EXAMPLE 4: Identification of candidate breast cancer genes
by analysis of genomic DNA sequence

Selected minimal overlapping BAC and PAC clones from
the physical map contig were sequenced in order to aid in
5   the identification of candidate breast cancer genes. DNA
was prepared from selected clones using a large scale DNA
isolation kit (Qiagen). Approximately 25-50 ug of DNA was
then sheared by nebulisation (10psi for 45 seconds) and
blunt ended using standard methodologies (Sambrook et al.,
10   1989). Samples were then run on an agarose gel in order to
isolate DNA in the 2-4 Kb size range. These fragments were
cleaned from the agarose using QIAquick columns (Qiagen),
ligated into puc18 and used to transform competent DH10B
or DH5a *E. coli* cells. DNA was isolated from transformed
15   clones and was sequenced using vector specific primers on
an ABI377 sequencer.

Analysis of genomic sequence was performed using
PHRED, PHRAP and GAP4 software on a SUN workstation. To
assist in the generation of large contigs of genomic
20   sequence, information present in the high-throughput
genomic sequence (htgs) database at NCBI was incorporated
into the assembly phase of the sequence analysis. The
resultant genomic sequence contigs were masked for repeats
and analysed using the BLAST algorithm (Altschul et al.,
25   1997) to identify nucleotide and protein homology to
sequences in the GenBank non-redundant and EST databases
at NCBI. The genomic sequence was also analysed for
predicted gene structure using the GENSCAN program and
specific screening of the mouse EST dataset was utilised
30   to identify potential human orthologues that have poor
representation in the human EST dataset.

Following the identification of homologous EST
sequences, *in silico* cDNA walking experiments were
initiated through further dbEST database screening. This
35   was to identify overlapping cDNA sequences present in
dbEST that would allow extension of the originally
identified partial gene sequence. Overlapping EST

sequences were assembled using the DNAStar LaserGene sequence assembly software. Homologous IMAGE cDNA clones in some instances were also purchased and sequenced. These longer stretches of sequence were then compared to known

5 genes by nucleotide and amino acid sequence comparisons using the above procedures.

From *in silico* analysis of the dbEST database at NCBI using all genomic sequence obtained for the 16q24.3 critical LOH region, a total of 55 gene fragments or gene

10 "signatures" were identified. In the majority of cases each novel gene fragment was represented by a distinct UniGene cluster composed of one or a number of overlapping cDNA clones. The majority of these UniGene clusters appeared to represent the 3' untranslated regions of their

15 representative gene as their sequence was continuous with the genomic sequence and further *in silico* manipulation failed to identify open reading frames representing amino acid coding regions.

As well as the 55 gene signatures that were

20 identified in the 16q24.3 region analysed, a total of 48 partial or full-length genes were also present based on *in silico* analysis of the genomic DNA generated.

Those sequences that are expressed in the breast were considered to be the most likely candidate breast cancer

25 genes. Those genes whose function could implicate it in the tumourigenic process, as predicted from homology searches with known proteins, were treated with the highest priority. Further evidence that a particular candidate is the responsible gene comes from the

30 identification of defective alleles of the gene in affected individuals or from analysis of the expression levels of a particular candidate gene in breast cancer samples compared with normal control tissues.

35 **EXAMPLE 5: Examination of the expression level of breast cancer gene candidates**

To investigate a potential role in breast cancer of

the genes identified from the 16q24.3 LOH region, the
level of expression of these genes, in a set of breast
cancer cell lines, was compared with their expression in
normal tissue controls. Differential expression (observed
5   as either a down-regulation or up-regulation of gene
expression) of a particular gene in a cancer cell line
compared to normal controls provides evidence that the
gene may be implicated in the cancer. The differential
expression may be due to point mutations in the gene,
10   which can decrease the stability of the mRNA of the gene
(viewed as a down-regulation of expression) or may lead to
enhanced expression of the gene (viewed as up-regulation
of expression). In addition, epigenetic mechanisms such as
abnormal promoter methylation may have the effect of
15   switching off gene expression which will also be observed
as a down-regulation in expression of the associated gene.
Recent studies have shown that this latter mechanism has
been responsible for the inactivation of other tumour
suppressor genes such as RB1 (Ohtani-Fujita et al., 1997),
20   VHL (Prowse et al., 1997), MLH1 (Herman et al., 1998) and
BRCA1 (Esteller et al., 2000).

        To detect the level of expression of the genes
identified in the 16q24.3 region in cancer samples
compared with normal controls, quantitative RT-PCR using
25   individual gene specific primers was done. This initially
involved the isolation of RNA from cancer cell lines along
with appropriate cell line and normal tissue controls.

Breast Cancer Cell Lines and RNA Extraction
30   Breast cancer cell lines were purchased from ATCC
(USA) and grown in the recommended tissue culture medium.
The breast cancer cell lines that were chosen for RT-PCR
analysis demonstrated homozygosity for a set of markers
mapping to chromosome 16q indicating potential LOH for
35   this chromosomal arm. Cells were harvested from confluent
cultures and total RNA was extracted using the RNAeasy kit
(Qiagen) or the TRIzol™ reagent (Gibco BRL) according to

manufacturers recommendations. PolyA$^+$ mRNA was subsequently isolated from all sources using the Oligotex bead system (Qiagen) according to recommended procedures. Total RNA derived from 21 human tissues (18 adult and 3 fetal) was

5   purchased commercially (Clontech, Stratagene, Ambion). DNA contamination was removed from all RNA preparations using DNAfree (Ambion) according to manufacturers protocols.

Reverse Transcription

10   Total RNA and PolyA$^+$ mRNA was primed with oligo-dT primers and reverse transcribed using the Omniscript RT kit (Qiagen) according to manufacturers conditions or using Superscript^{TM} RNaseH$^-$ reverse transcriptase (Gibco BRL). In the latter method, 1 ug of total RNA sample was

15   mixed with 500 ng of oligo (dT)$_{16}$ and made up to a volume of 10 ul with DEPC treated water. Following a 10 minute incubation at 70°C, 4 ul of 5X first strand buffer, 2 ul of 0.1 M DTT, 1 ul of 10 mM dNTP, 20 units of RNAsin^{TM} (Promega) and 100 units of Superscript reverse

20   transcriptase were added and the reaction incubated at 42°C for 2 hours. Reactions were terminated at 95°C for 5 minutes and cDNA:RNA hybrids were removed from samples by addition of 2 units of RNase H (Promega) and incubation at 37°C for 30 minutes. Control reactions were included for

25   each RNA template, which omitted reverse transcriptase from the cDNA synthesis step. This was to determine the presence of any genomic DNA contamination in the RNA samples. All samples were stored at -20°C.

30 cDNA Normalisation

       Internal standard curve amplicons were generated from a mixed pool of normal tissue cDNA using the HotStarTaq^{TM} DNA Polymerase kit (Qiagen). A reaction mix sufficient to generate >1 ug of amplicon cDNA contained 10 ul of 10× PCR

35   buffer (containing 15 mM MgCl$_2$), 2 ul of 10 mM dNTP mix, 0.5 uM of each primer, 0.5 ul of 2.5 units HotStarTaq polymerase (Qiagen), 100 ng of cDNA template and DEPC

- 44 -

treated water to 100 ul. Amplification cycling was performed as follows: 94°C for 10 minutes followed by 35 cycles at 93°C for 20 seconds, 60°C for 30 seconds and 70°C for 30 seconds with a final extension at 72°C for 4

5   minutes. Amplicons were purified using the QIAquick gel extraction kit (Qiagen) according to manufacturers conditions and concentrations were measured at A_{260}. Purified amplicons were serially diluted 10-fold from 10 ng/ul to 1 fg/ul. These dilutions served as internal

10 standards of known concentration for real-time analysis of each gene specific amplicon as described below.

Real-time PCR

    All cDNA templates were amplified using the SYBR

15 Green I PCR Master Mix kit (PE Biosystems, USA). Primer sets for the amplification of each gene were selected using the Lasergene Primer Select™ software (DNASTAR). PCR reactions were in a volume of 25 ul and included 12.5 ul of SYBR Green I PCR Master mix, 0.5 uM of each primer, 2

20 ul normalised cDNA template (see below) and 9.5 ul of water. Real-time PCR analysis was performed using the Rotor-Gene™2000 (Corbett Research, AUS) with the following amplification cycling conditions: 94°C for 10 minutes followed by 45 cycles of 93°C for 20 sec, 60°C for 30 sec

25 and 70°C for 30 sec. Fluorescence data was acquired at 510 nm during the 72°C extension phase. Melt curve analyses were performed with an initial 99-50°C cycling followed by fluorescence monitoring during heating at 0.2°C/second to 99°C. Prior to real-time quantification, product size and

30 specificity was confirmed by ethidium bromide staining of 2.5% agarose gels following electrophoresis of completed PCRs.

Real-time PCR Quantification

35   Quantification analyses were performed on the Rotor-Gene™ DNA sample analysis system (Version 4.2, Build 96). Standard curves were generated by amplifying 10-fold

serial dilutions (1 ul of 10 ρg/ul down to 1 ul of 1 fg/ul
in triplicate) of the internal standard amplicon during
real-time PCR of gene specific amplicons from normal
tissues and breast cancer cell lines. Internal standard

5   amplicon concentrations were arbitrarily set to 1.0e+12
copies for 10 ρg standards to 1.0e+08 copies for 1 fg
standards. C_T (cycle threshold) coefficients of variation
for all internal standard dilutions averaged 2% between
triplicate samples within the same and different runs. The

10 Rotor-Gene™ quantification software generated a line of
best-fit at the parameter C_T and determined unknown normal
tissue and breast cancer cell line amplicon copy numbers
by interpolating the noise-band intercept of each amplicon
against the internal standards with known copy numbers.

15

Normalization and relative expression of data

Using the expression value of each gene in normal
breast tissue as a baseline, the relative fold-difference
between the cell line exhibiting the highest expression

20 and the cell line exhibiting the lowest expression was
calculated. This value was termed the "Relative Fold
Variability Index" (RFVI).

In order to establish an RFVI baseline range, five
house-keeping genes were first examined. These included

25 Cyclophilin, APRT, RNA Polymerase subunit II, ATP synthase
and GAPDH. This baseline range reflects mRNA expression
differences that are due to normal population variations
or experimental reproducibility.

The degree of variation in mRNA expression levels for

30 the housekeeping genes was relatively uniform between the
normal tissues and cancer cell lines examined. Three-way
combinations for normalization between Cyclophilin, RNA
polymerase II subunit and APRT demonstrated a mean 7-fold
and maximum 50-fold variance in mRNA expression level

35 between samples. The significance of variable mRNA
expression levels within a gene of interest may therefore
reasonably be evaluated based on these normalization

results. A predicted aberrant alteration in gene of
interest mRNA copy number of >50 fold in breast cancer
cell lines relative to a 'baseline' normal breast
expression level was therefore considered to be
5 significantly abnormal.

Following establishment of a baseline RFVI value of
50, the RFVI was determined for the SYK and p16INK4a
genes. These represent known tumour suppressor genes that
have been previously shown to exhibit aberrant expression
10 in breast cancer cells (Coopman et al., 2000; Bisogna et
al., 2001). Both genes had significant RFVI values with
the observed RFVI for SYK being almost 10 times greater
than the baseline range (baseline=50; SYK=460).

Results from the examination of the RFVI values for
15 the genes identified in the 16q24.3 region are shown in
Figure 2. A total of 7 genes were identified to have an
RFVI of greater than 50 indicating significant
differential expression in breast cancer cell lines
compared with normal tissue controls. A description of
20 these genes is summarised in Table 1 and their sequences
are represented by the SEQ ID Numbers: 1-11.

The data generated from the expression studies is
also represented in Figure 3. In this instance, the
percent fold change in expression of each gene across the
25 entire panel of breast cancer cell lines examined is
displayed. Results indicate that the differential
expression in 6 of the 7 genes described above (BNO227,
BNO8, BNO205, BNO221, BNO225, BNO226) was due to a down-
regulation in gene expression, while the remaining gene
30 (BNO223) showed an up-regulation in gene expression
specific for a number of breast cancer cell lines.

BNO227 was identified to display the most significant
differential expression and corresponds to the CYBA gene.
CYBA associates with CYBB to form cytochrome b-558, which
35 is the membrane component of NADPH oxidase and functions
as the final electron transporter in the oxidation of
NADPH, resulting in the generation of Reactive Oxygen

- 47 -

Species (ROS) such as O_2 and H_2O_2. The levels of ROS appear to be critical in the regulation of a number of genes involved in diverse pathways delineating transcription, proliferation and apoptosis (Burdon et al, 1996; Arnold et
5   al, 2001; Jacobson, 1996). This regulatory role is thought to encompass influences on gene expression and protein function (Burdon, 1995). It has been suggested that alteration of protein function is effected by modifications on redox-sensitive amino acids such as
10 cystein and histidine. The role of CYBA as a NADPH oxidase in the microbicidal function of phagocytes has been studied extensively. Mutations in this gene are causative to the chronic granulomatous disease characterized by recurrent bacterial and fungal infections (Rae et al,
15 2000). More recently, CYBA has been implicated in oxidases involved in epithelial and muscle cell gene regulation and function with demonstrated implications in atherosclerosis (Sorescu et al, 2001). Hitherto, there has been no direct demonstration of a CYBA NADPH oxidase role in breast
20 cancer, however, several studies have shown the involvement of reactive oxygen species (O2*-, hypochlorite, hydroxyl radical, hydrogen peroxide) in carcinogenesis and tumour progression (Gupta et al, 2001; Brown and Bicknell, 2001). The relative concentration of
25 ROS is critical to their function. ROS levels are under a tight regulatory control involving the interplay of NADPH oxidases and antioxidant ROS scavengers (Griendling and Ushio-Fukai, 2000). Any disruption to these control mechanisms is likely to result in aberrant cell behaviour
30 as that seen in cancer. Consequently, the expression profiles of the molecules involved in ROS production and/or removal are of central importance. This example demonstrates that CYBA is expressed in normal breast tissue and in breast cancer cell lines. In addition, we
35 have found that this gene is differentially expressed, with some cell lines expressing very low levels of this gene. This finding implicates CYBA as a potential tumour

- 48 -

suppressor and suggests the possible involvement in
carcinogenesis of the other membrane subunit of
cytochrome-b558 as well as the cytoplasmic components of
NADPH oxidase.

5

EXAMPLE 6: Analysis of tumours and cell lines for breast
cancer gene mutations

     Any one of the genes that have been shown to be
differentially expressed in this study can be screened by
10   single strand conformation polymorphism (SSCP) analysis in
DNA isolated from tumours which display restricted LOH for
the 16q24.3 region. This can be done to identify those
samples where mutations in the gene are causative for the
cancer rather than disregulation of gene expression being
15   the causative factor. In this instance DNA isolated from
series 1 and series 2 tumours can be used. A number of
breast cancer cell lines, or cell lines from other cancer
types, may also be screened. Likewise, tissues from other
cancer types can be screened by SSCP for disease causing
20   mutations. Cell lines can be purchased from ATCC, grown
according to manufacturers conditions, and DNA isolated
from cultured cells using standard protocols (Wyman and
White, 1980; Sambrook et al., 1989).

     To perform mutation analysis of the candidate breast
25   cancer genes using the SSCP technique, a number of
variations can be employed. For example, breast cancer
gene exons can be amplified by PCR using flanking intronic
primers, which are labeled at their 5' ends with HEX.
Typical PCR reactions are performed in 96-well plates in a
30   volume of 10 ul using 30 ng of template DNA. Cycling
conditions involve an initial denaturation step at 94°C for
3 minutes followed by 35 cycles of 94°C for 30 seconds,
60°C for $1^{1/2}$ minutes and 72°C for $1^{1/2}$ minutes. A final
extension step of 72°C for 10 minutes follows. Twenty ul of
35   loading dye comprising 50% (v/v) formamide, 12.5 mM EDTA
and 0.02% (w/v) bromophenol blue is added to completed
reactions which are subsequently run on 4% polyacrylamide

gels and analysed on the GelScan 2000 system (Corbett Research, AUS) according to manufacturers specifications.

Those samples that display a bandshift compared with normal controls are considered to have a different nucleotide composition in the amplicon being analysed compared to that of normal controls. The amplicon can be sequenced in this sample and compared to wild-type sequence to determine the nucleotide differences. Any base changes that are present in a tumour sample but not present in the corresponding normal control sample from the same individual or in other normal individuals most likely represents a deleterious mutation. This is further confirmed if the base change also leads to an amino acid change or the generation of a truncated form of the protein.

EXAMPLE 7: Analysis of the breast cancer genes

The following methods are used to determine the structure and function of any one of the breast cancer genes.

Biological studies

Mammalian expression vectors containing breast cancer gene cDNA can be transfected into breast, prostate or other carcinoma cell lines that have lesions in the gene. Phenotypic reversion in cultures (eg cell morphology, growth of transformants in soft-agar, growth rate) and in animals (eg tumourigenicity in nude mice) is examined. These studies can utilise wild-type or mutant forms of the breast cancer genes. Deletion and missense mutants of these genes can be constructed by *in vitro* mutagenesis.

Molecular biological studies

The ability of any one of the breast cancer proteins to bind known and unknown proteins can be examined. These proteins may give an insight as to the biological pathways in which the breast cancer proteins participate. In turn,

proteins within these pathways may provide suitable targets for therapeutic applications such as gene therapy, screening for small molecule interactors, as well as antisense and antibody-based therapies directed at these
5   interactors.

Procedures such as the yeast two-hybrid system are used to discover and identify any functional partners. The principle behind the yeast two-hybrid procedure is that many eukaryotic transcriptional activators, including
10   those in yeast, consist of two discrete modular domains. The first is a DNA-binding domain that binds to a specific promoter sequence and the second is an activation domain that directs the RNA polymerase II complex to transcribe the gene downstream of the DNA binding site. Both domains
15   are required for transcriptional activation as neither domain can activate transcription on its own. In the yeast two-hybrid procedure, the gene of interest or parts thereof (BAIT), is cloned in such a way that it is expressed as a fusion to a peptide that has a DNA binding
20   domain. A second gene, or number of genes, such as those from a cDNA library (TARGET), is cloned so that it is expressed as a fusion to an activation domain. Interaction of the protein of interest with its binding partner brings the DNA-binding peptide together with the activation
25   domain and initiates transcription of the reporter genes. The first reporter gene will select for yeast cells that contain interacting proteins (this reporter is usually a nutritional gene required for growth on selective media). The second reporter is used for confirmation and while
30   being expressed in response to interacting proteins it is usually not required for growth.

Structural studies

Breast cancer recombinant proteins can be produced in
35   bacterial, yeast, insect and/or mammalian cells and used in crystallographical and NMR studies. Together with

- 51 -

molecular modeling of the proteins, structure-driven drug
design can be facilitated.

EXAMPLE 8: Generation of polyclonal antibodies against the
5   breast cancer proteins
        The knowledge of the nucleotide and amino acid
sequence of the breast cancer genes and associated
proteins allows for the production of antibodies, which
selectively bind to these proteins or fragments thereof.
10 Following the identification of mutations in these breast
cancer genes, antibodies can also be made to selectively
bind and distinguish mutant from normal protein.
Antibodies specific for mutagenised epitopes are
especially useful in cell culture assays to screen for
15 malignant cells at different stages of malignant
development. These antibodies may also be used to screen
malignant cells, which have been treated with
pharmaceutical agents to evaluate the therapeutic
potential of the agent.
20   To prepare polyclonal antibodies, short peptides can
be designed homologous to any one of the breast cancer
amino acid sequences. Such peptides are typically 10 to 15
amino acids in length. These peptides should be designed
in regions of least homology to the mouse orthologue to
25 avoid cross species interactions in further down-stream
experiments such as monoclonal antibody production.
Synthetic peptides can then be conjugated to biotin
(Sulfo-NHS-LC Biotin) using standard protocols supplied
with commercially available kits such as the PIERCE™ kit
30 (PIERCE). Biotinylated peptides are subsequently complexed
with avidin in solution and for each peptide complex, 2
rabbits are immunized with 4 doses of antigen (200 ug per
dose) in intervals of three weeks between doses. The
initial dose is mixed with Freund's Complete adjuvant
35 while subsequent doses are combined with Freund's Immuno-
adjuvant. After completion of the immunization, rabbits
are test bled and reactivity of sera assayed by dot blot

- 52 -

with serial dilutions of the original peptides. If rabbits
show significant reactivity compared with pre-immune sera,
they are then sacrificed and the blood collected such that
immune sera can separated for further experiments.

5

EXAMPLE 9: Generation of monoclonal antibodies specific
for the breast cancer proteins

Monoclonal antibodies can be prepared for any one of
the breast cancer proteins in the following manner.
10 Immunogen comprising an intact breast cancer protein or
peptide (wild type or mutant) is injected in Freund's
adjuvant into mice with each mouse receiving four
injections of 10 to 100 ug of immunogen. After the fourth
injection blood samples taken from the mice are examined
15 for the presence of antibody to the immunogen. Immune mice
are sacrificed, their spleens removed and single cell
suspensions are prepared (Harlow and Lane, 1988). The
spleen cells serve as a source of lymphocytes, which are
then fused with a permanently growing myeloma partner cell
20 (Kohler and Milstein, 1975). Cells are plated at a density
of $2X10^5$ cells/well in 96 well plates and individual wells
are examined for growth. These wells are then tested for
the presence of specific antibodies by ELISA or RIA using
wild type or mutant breast cancer target protein. Cells in
25 positive wells are expanded and subcloned to establish and
confirm monoclonality. Clones with the desired specificity
are expanded and grown as ascites in mice followed by
purification using affinity chromatography using Protein A
Sepharose, ion-exchange chromatography or variations and
30 combinations of these techniques.

Industrial Applicability

The DNA sequences of the present invention are useful
in the diagnosis of cancer, or a pre-disposition thereto
35 and, where they are not full-length gene sequences, they
may be used to identify full-length genes involved in

carcinogenesis.   Methods   of   treatment   of   cancer   and
methods of screening for drugs are also made available.

5

TABLE 1

Differentially Expressed Breast Cancer Genes Identified at Chromosome 16q24.3

Gene	Associated UniGene Cluster	Description	SEQ ID Numbers
BNO227[1]	Hs.68877	Cytochrome b-245, alpha polypeptide (CYBA)	1, 2
BNO8[1]	Hs.189419	ESTs	3
BNO205[1]	Hs.174074	ESTs	4
BNO221[1]	None	EST	5
BNO225[1]	Hs.148090	Cadherin 15 (CDH15)	6, 7
BNO226[1]	Hs.109	Dipeptidase 1 (DPEP1)	8, 9
BNO223[2]	Hs.184601	Solute carrier family 7 (cationic amino acid transporter), member 5 (SLC7A5)	10, 11

Note: [1] Down-regulation of expression in breast cancer cell lines compared to normal tissue controls. [2] Up-regulation of expression in breast cancer cell lines compared to normal tissue controls.

- 55 -

References

References cited herein are listed on the following pages, and are incorporated herein by this reference.

5   Altschul, SF. et al. (1997). *Nucleic Acids Res.* 25: 3389-3402.

Arnold, RS. et al. (2001). *Proc. Natl. Acad. Sci. USA* 98: 5550-5555.

Bisogna, M. et al. (2001). *Cancer Genet. Cytogenet.* 125:
10   131-138.

Brenner, AJ. and Aldaz CM. (1995). *Cancer Res.* 55: 2892-2895.

Brown, NS. and Bicknell, R. (2001). *Breast Cancer Res.* 3: 323-327.

15 Burdon, RH. (1995). *Free Radic. Biol. Med.* 18: 775-794.

Burdon, RH. et al. (1996). *Free Radic. Res.* 24: 81-93.

Callen, DF. et al. (1995). *Genomics* 29: 503-511.

Chen, T. et al. (1996). *Cancer Res.* 56: 5605-5609.

Cleton-Jansen, A-M. et al. (1995). *Br. J. Cancer* 72: 1241-
20   1244.

Cole, SP. et al. (1984). *Mol. Cell Biol.* 62: 109-120.

Coopman, PJ. et al. (2000). *Nature* 406: 742-747.

Cote, RJ. et al. (1983). *Proc. Natl. Acad. Sci. USA* 80: 2026-2030.

25 Culver, K. (1996). *Gene Therapy : A Primer for Physicians.* Second Edition. (Mary Ann Liebert).

Devilee, P. et al. (1991). *Oncogene* 6: 1705-1711.

Devilee, P. et al. (1994). *Genes Chrom. Cancer* 11: 71-78.

Devilee, P. and Cornelisse, CJ. (1994). *Biochimica et*
30   *Biophysica Acta* 1198: 113-130.

Doggett, NA. et al. (1995). *Nature* 377 Suppl: 335-365.

Elston, CW. and Ellis, IO. (1990). *Histopathology* 16: 109-118.

Esteller, M. et al. (2000). *J. Natl. Cancer Inst.* 92: 564-
35   569.

Fearon, ER. and Vogelstein, B. (1990). *Cell* 61: 759-767.

- 56 -

Friedman, T. (1991). In *Therapy for Genetic Diseases*. T Friedman (Ed). Oxford University Press. pp 105-121.

Futreal, PA. et al. (1994). *Science* 266: 120-122.

Gecz, J. et al. (1997). *Genomics* 44: 201-213.

Goldman, CK. et al. (1997). *Nature Biotechnology* 15: 462-466.

Griendling, KK. and Oshio-Fukai, M. (2000). *Regulatory Peptides* 91: 21-27.

Groden, J. *et al.* (1991). *Cell* 66: 589-600.

Gupta, A. et al. (2001). *Cancer Lett.* 173: 115-125.

Hall, JM. et al. (1990). *Science* 250: 1684-1689.

Harlow, E. and Lane, D. (1988). *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

Heller, RA. et al. (1997). *Proc. Natl. Acad. Sci. USA* 94: 2150-2155.

Herman, JG. et al. (1998). *Proc. Natl. Acad. Sci. USA* 95: 6870-6875.

Huse, WD. et al. (1989). *Science* 246: 1275-1281.

Kohler, G. and Milstein, C. (1975). *Nature* 256: 495-497.

Kozbor, D. et al. (1985). *J. Immunol. Methods* 81:31-42.

Longmire, JL. et al. (1993). *GATA* 10: 69-76.

McCormick, MK. et al. (1993). *Proc. Natl. Acad. Sci. USA* 90: 1063-1067.

Miki, Y. et al. (1994). *Science* 266: 66-71.

Miki, Y. et al. (1996). *Nature Genet.* 13: 245-247.

Ohtani-Fujita, N. et al. (1997). *Cancer Genet. Cytogenet.* 98:43-49.

Prowse, AH. et al. (1997). *Am. J. Hum. Genet.* 60:765-771.

Radford, DM. et al. (1995). *Cancer Res.* 55: 3399-3405.

Rae, J. et al. (2000). *Blood* 96: 1106-1112.

Riethman, HC. et al. (1989). *Proc. Natl. Acad. Sci. USA* 86: 6240-6244.

Saito, H. et al. (1993). *Cancer Res.* 53: 3382-3385.

Sambrook, J. et al. (1989). *Molecular cloning: a laboratory manual*. Second Edition. (Cold Spring Harbour Laboratory Press, New York).

- 57 -

Scharf, D. et al. (1994). *Results Probl. Cell Differ.* 20: 125-162.

Schena, M. et al. (1996). *Proc. Natl. Acad. Sci. USA* 93: 10614-10619.

5   Soares, MB. et al. (1994). *Proc. Natl. Acad. Sci. USA* 91: 9228-9232.

Sorescu, D. et al. (2001). *Trends Cardiovasc. Med.* 11: 124-131.

Weber, JL. and May, PE. (1989). *Am. J. Hum. Genet.* 44: 10   388-396.

Whitmore, SA. et al. (1994). *Genomics* 20: 169-175.

Whitmore, SA. et al. (1998). *Genomics* 50: 1-8.

WHO. (1981). *Histological Typing of Breast Tumours.* Second Edition. (Geneva).

15   Wooster, R. et al. (1995). *Nature* 378: 789-791.

Wooster, R. et al. (1994). *Science* 265: 2088-2090.

Wyman, A. and White, R. (1980). *Proc. Natl. Acad. Sci. USA* 77: 6754-6758.

20

Claims

1.   A method for the diagnosis of cancer, or a predisposition thereto, in a patient, comprising the steps of:

      (1)   establishing the level of expression of a gene selected from the group consisting of cytochrome B-245, alpha polypeptide (CYBA), cadherin 15 (CDH15), dipeptidase 1 (DPEP1), solute carrier family 7, member 5 (SLC7A5), BNO8, BNO205 and BNO221; and

      (2)   comparing expression of the gene to a baseline established from expression in normal tissue controls;

      wherein substantial variance from the baseline indicates that the patient has, or is susceptible to, cancer.

2.   A method as claimed in claim 1 wherein down-regulation of CYBA, CDH15, DPEP1, BNO8, BNO205 or BNO221 is observed.

3.   A method as claimed in claim 1 wherein up-regulation of SLC7A5 is observed.

4.   A method for determining whether a human tissue is predisposed to a neo-plastic transformation, comprising determining whether in a cell from the tissue a nucleic acid molecule selected from the group consisting of CYBA, CDH15, DPEP1, BNO8, BNO205 or BNO221 is absent, present in a mutant form or down-regulated through epigenetic mechanisms.

5.   A method according to claim 4 wherein the human tissue is human breast tissue.

- 59 -

6.   A method as claimed in claim 4 comprising determining whether the encoded polypeptide is absent or expressed at reduced levels.

5   7.   A method as claimed in claim 6 comprising contacting the cell with an antibody for binding the peptide under conditions which permit the antibody to bind the peptide if it is present.

10   8.   A method for determining whether a human tissue has undergone a neoplastic transformation, comprising, determining whether in a cell from the tissue the SLC7A5 gene is up-regulated and/or present in mutant form.

15   9.   A method as claimed in claim 8 wherein the human tissue is human breast tissue.

10.   A method as claimed in claim 8 comprising determining whether the encoded polypeptide is over-expressed and/or a
20   mutant peptide is expressed.

11.   A method as claimed in claim 10 comprising contacting the cell with an antibody specific for the mutant peptide under conditions which permit the antibody to bind to the
25   mutant peptide, if present.

12.   A mutant form of CYBA, CDH15, DPEP1, BNO8, BNO205 or BNO221 in which the tumour suppressor activity of the gene is compromised.
30
13.   A mutant form of SLC7A5 which renders it active as an oncogene.

14.   A polypeptide encoded by a mutant form of a gene as
35   defined in claim 12.

15. A polypeptide encoded by a mutant form of a gene as defined in claim 13.

16. An antibody to a mutant form of a gene as defined in claim 12.

17. An antibody to a mutant form of a gene as defined in claim 13.

18.   Use of a DNA molecule selected from the group consisting of DNA molecules having the sequence set forth in SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5 to identify and/or obtain full-length human genes involved in the tumourigenic process.

19.   Use as claimed in claim 18 wherein a full-length human gene is identified by hybridisation with one or more of said nucleotide molecules.

20.   Use as claimed in either one of claims 18 or 19 wherein additional sequence is obtained using inverse PCR, restriction site PCR, PCR walking techniques or RACE.

21.   Use as claimed in claim 18 wherein PCR probes are designed based upon the nucleotide sequence of said DNA molecule.

22.   Use as claimed in claim 21 wherein the probes comprise 5 to 20 nucleotides from said DNA molecule.

23.   Use as claimed in claim 22 wherein the probes are labelled with radionuclides or enzymatic labels.

24.   Use as claimed in any one of claims 18 to 23 wherein the full-length human gene is a tumour suppressor gene.

25.   A method of identifying a gene involved in the
tumourigenic process, comprising the steps of:

          (1)   providing a nucleic acid probe derived from
or comprising a DNA molecule selected from the group
consisting of DNA molecules having the sequence set forth
in SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5, or a fragment
thereof;

          (2)   assembling   a   hybridisation   mixture
containing said nucleic acid probe in single stranded form
and a plurality of single stranded nucleic acid molecules
derived from a region of loss of heterozygosity at
16q24.3; and

          (3)   isolating a double stranded hybrid of said
nucleic acid probe and a complementary nucleic acid
molecule derived from a region of loss of heterozygosity
at 16q24.3, the DNA molecule from which the complementary
nucleic acid molecule is derived representing at least a
portion of a candidate gene.

26.   A gene when identified by the method of claim 25.

27.   A gene incorporating the sequence set forth in
any of SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5.

28.   A gene when identified through the use as claimed
in any one of claims 18 to 24.

29.   A gene located at 16q24.3 which hybridises under
stringent conditions to one or more nucleotide molecules
selected from the group consisting of nucleotide molecules
derived from DNA molecules having the sequence set forth
in SEQ ID NO:3, SEQ ID NO:4 and SEQ ID NO:5.

30.   A gene as claimed in any one of claims 26 to 29
which is a tumour suppressor gene.

31.   A polypeptide encoded by a gene as claimed in any

one of claims 26 to 30.

32.   An antibody to a polypeptide as defined in claim
31.

33.   An antibody to a polypeptide having an amino acid
sequence set forth in SEQ ID NO:2, SEQ ID NO:7, SEQ ID
NO:9 and SEQ ID NO:11.

34.   Use of a gene as claimed in claims 26 to 30 in
the diagnosis of cancer or in establishing the prognosis
of a patient diagnosed with cancer.

35.   Use of a nucleotide molecule having the
nucleotide sequence set forth in SEQ ID NO:1, SEQ ID NO:3,
SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:8 and SEQ
ID NO:10 in the diagnosis of cancer or in establishing the
prognosis of a patient diagnosed with cancer.

36.   Use as claimed in claims 34 and 35 wherein the
cancer is breast cancer.

37.   Use of a polypeptide as defined in claim 31 in
the diagnosis of cancer or in establishing the prognosis
of a patient diagnosed with cancer.

38.   Use of a polypeptide having the amino acid
sequence set forth in SEQ ID NO:2, SEQ ID NO:7, SEQ ID
NO:9 and SEQ ID NO:11 in the diagnosis of cancer or in
establishing the prognosis of a patient diagnosed with
cancer.

39.   Use as claimed in claims 37 and 38 wherein the
cancer is breast cancer.

40.   Use of an antibody as claimed in claim 32 or 33
in the diagnosis of cancer or in establishing the

- 63 -

prognosis of a patient diagnosed with cancer.

41.   Use as claimed in claim 40 wherein the cancer is breast cancer.

42.   A microarray comprising oligonucleotides or longer fragments derived from any one or more of the genes CYBA, CDH15, DPEP1, SLC7A5 or a gene as claimed in any one of claims 26 to 30.

43.   A method of treating or inhibiting cancer in a patient in need of such treatment, said method comprising administering to said patient a vector capable of expressing CYBA, CDH15, DPEP1 or a gene as claimed in any one of claims 26 to 30.

44.   A method of treating or inhibiting cancer in a patient in need of such treatment, said method comprising administering to said patient a compound which increases expression or activity of CYBA, CDH15, DPEP1 or a gene as claimed in any one of claims 26 to 30.

45.   A method of treating or inhibiting cancer in a patient in need of such treatment, said method comprising administering to said patient an antagonist of SLC7A5.

46.   A method as claimed in claim 45 wherein the antagonist of SLC7A5 is an antibody thereto, or an antisense thereto, or a nucleic acid complementary in sequence to at least a portion of SLC7A5.

47.   A method as claimed in any of claims 43 to 46 wherein the cancer is breast cancer.

48.   Use of a gene as claimed in any of claims 26 to 30 in the treatment of cancer through gene therapy.

49.   Use of a nucleotide molecule having the nucleotide sequence set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10 in the treatment of cancer through gene therapy.

50.   Use of a polypeptide as defined in claim 31 in the treatment of cancer.

51.   Use of a polypeptide having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:7, SEQ ID NO:9 and SEQ ID NO:11 in the treatment of cancer.

52.   Use of an antibody as claimed in either one of claims 32 or 33 in the treatment of cancer.

53.   Use as claimed in any one of claims 48 to 52 wherein the cancer is breast cancer.

54.   A genetically modified non-human animal in which a gene selected from the group consisting of CYBA, CDH15, DPEP1, BNO8, BNO205, BNO221, or a gene as claimed in any one of claims 26 to 30 has been inactivated by knockout deletion.

55.   A genetically modified non-human animal in which the SLC7A5 gene is over-expressed and/or expressed in mutant form as an oncogene.

56.   A genetically modified animal as claimed in either of claims 54 or 55 wherein the animal is selected from the group consisting of rats, mice, hamsters, guineas pigs, rabbits, dogs, cats, goats, sheep, pigs and non-human primates such as monkeys and chimpanzees.

57.   A method of screening for candidate drugs which restore tumour suppressor activity, comprising the steps of:

- 65 -

     (1)   contacting a cell, in which the functioning of a gene selected from the group consisting of CYBA, CDH15, DPEP1, BNO8, BNO205, BNO221, or a gene as claimed in any one of claims 26 to 30 is compromised, with a candidate drug; and

     (2)   assaying for the amount of tumour suppressor activity in the cell.

58.   A method of screening for candidate drugs which reverse oncogene activity, comprising the steps of:

     (1)   contacting a cell in which SLC7A5 is over-expressed and/or expressed in a mutant form with a candidate drug; and

     (2)   assaying for the level of expression and/or activity of SLC7A5 in the cell.

59.   A method of screening for a drug capable of binding SLC7A5 and modulating its activity, said method comprising combining the SLC7A5 protein and a candidate drug and determining the binding of said candidate drug to the SLC7A5 peptide.

60.   Use of a drug identified by the method of any one of claims 57 to 58 in the treatment of cancer.

61.   Use of a drug identified by the method any one of claims 57 to 59 in the manufacture of a medicament for the treatment of cancer.

62.   Use as claimed in any one of claims 60 or 61 wherein the cancer is breast cancer.

63.   An anti-cancer drug when identified by the method of any one of claims 57, 58 or 59.

64.   A compound which increases the expression of or

activity of CYBA, CDH15, DPEP1, BNO8, BNO205, BNO221 or a gene as claimed in claim 26.

65.   An antagonist of SLC7A5.

66.   A cell comprising an expression vector capable of expressing CYBA, CDH15, DPEP1, BNO8, BNO205, BNO221 or a gene as claimed in any one of claims 26 to 30.

67.   A cell transformed with a nucleotide molecule having the nucleotide sequence set forth in SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10.

68.   A pharmaceutical composition comprising any one or more of the genes selected from the group consisting of CYBA, CDH15, DPEP1, SLC7A5, BNO8, BNO205, BNO221 or genes as claimed in any one of claims 26 to 30, or active fragments thereof, their expression products and antibodies to their expression products, or any one or more isolated nucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10, and an inert carrier.

69.   A pharmaceutical composition comprising any one or more of a polypeptide as claimed in claim 29 or 30 or having the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:7, SEQ ID NO:9 and SEQ ID NO:11, an antibody as claimed in claim 32 or 33, an expression vector according to claim 43, a compound according to claim 64, an anti-cancer drug according to claim 63, or an antagonist of SLC7A5 as claimed in claim 65, and an inert carrier.

Figure 1

Relative Fold Variability Index (RFVI)

Figure 2

% Fold change breast cancer cell line mRNA expression
relative to normal breast tissue
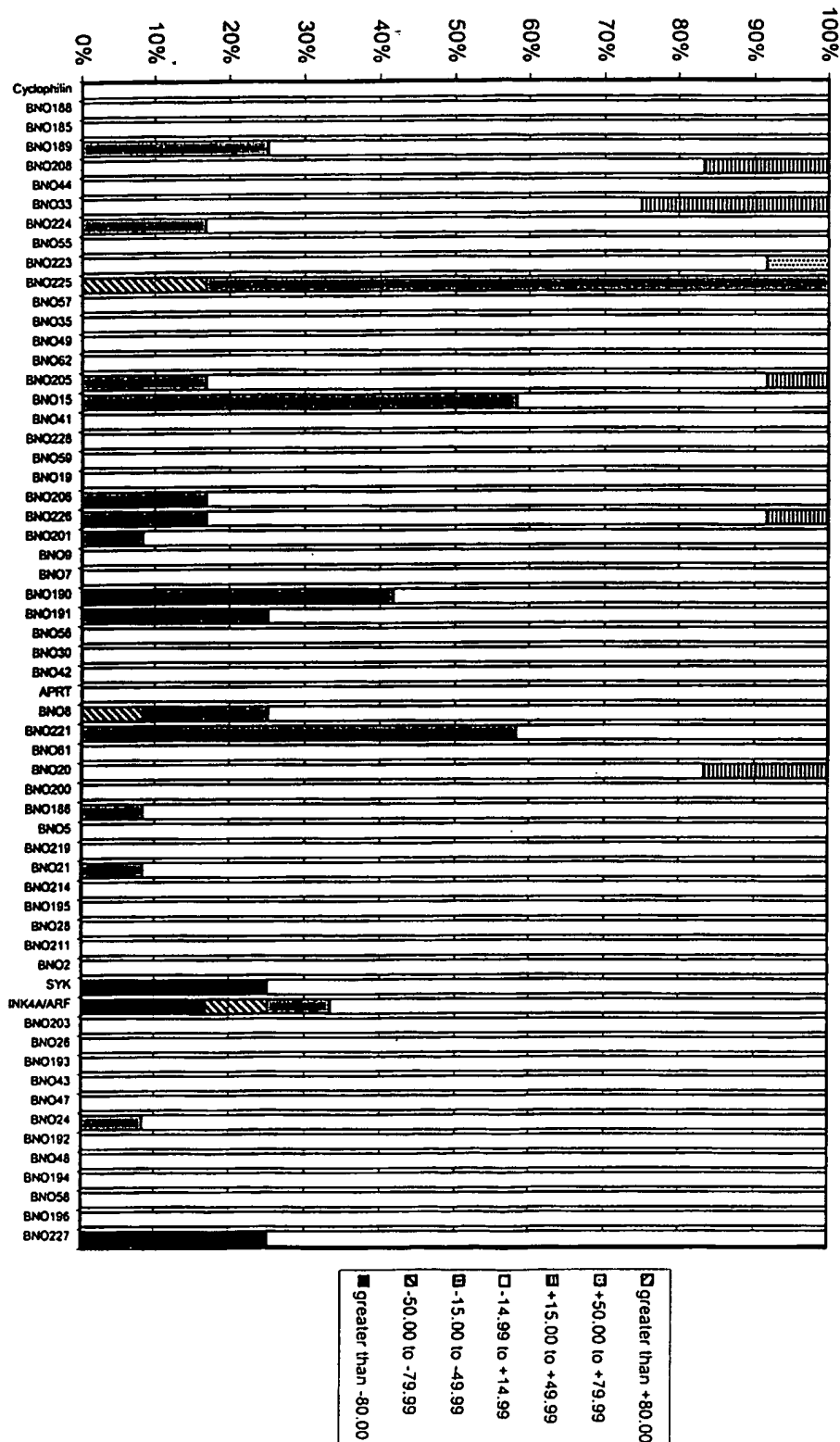
Figure 3

DiffExpr.ST25.txt
SEQUENCE LISTING

<110>   Bionomics Limited

<120>   P10

<130>   Breast Cancer - Differentially Expressed

<160>   11

<170>   PatentIn version 3.1

<210>   1
<211>   680
<212>   DNA
<213>   Homo sapiens

<400>   1
gcagtgtccc agccgggttc gtgtcgccat ggggcagatc gagtgggcca tgtgggccaa   60

cgagcaggcg ctggcgtccg gcctgatcct catcaccggg ggcatcgtgg ccacagctgg   120

gcgcttcacc cagtggtact ttggtgccta ctccattgtg gcgggcgtgt ttgtgtgcct   180

gctggagtac ccccggggga agaggaagaa gggctccacc atggagcgct ggggacagaa   240

gcacatgacc gccgtggtga agctgttcgg gccctttacc aggaattact atgttcgggc   300

cgtcctgcat ctcctgctct cggtgcccgc cggcttcctg ctggccacca tccttgggac   360

cgcctgcctg gccattgcga gcggcatcta cctactggcg gctgtgcgtg gcgagcagtg   420

gacgcccatc gagcccaagc cccgggagcg gccgcagatc ggaggcacca tcaagcagcc   480

gcccagcaac cccccgccgc ggcccccggc cgaggcccgc aagaagccca gcgaggagga   540

ggctgcggcg gcggcggggg gaccccccggg aggtccccag gtcaacccca tcccggtgac

DiffExpr.ST25.txt

600

cgacgaggtc gtgtgacctc gccccggacc tgccctccca ccaggtgcac ccacctgcaa

660

taaacgcagc gaaggccggg

680

<210> 2
<211> 195
<212> PRT
<213> Homo sapiens

<400> 2

Met Gly Gln Ile Glu Trp Ala Met Trp Ala Asn Glu Gln Ala Leu Ala
1   5   10   15

Ser Gly Leu Ile Leu Ile Thr Gly Gly Ile Val Ala Thr Ala Gly Arg
            20   25   30

Phe Thr Gln Trp Tyr Phe Gly Ala Tyr Ser Ile Val Ala Gly Val Phe
            35   40   45

Val Cys Leu Leu Glu Tyr Pro Arg Gly Lys Arg Lys Lys Gly Ser Thr
    50   55   60

Met Glu Arg Trp Gly Gln Lys His Met Thr Ala Val Val Lys Leu Phe
65   70   75   80

Gly Pro Phe Thr Arg Asn Tyr Tyr Val Arg Ala Val Leu His Leu Leu
                85   90   95

Leu Ser Val Pro Ala Gly Phe Leu Leu Ala Thr Ile Leu Gly Thr Ala
            100   105   110

Cys Leu Ala Ile Ala Ser Gly Ile Tyr Leu Leu Ala Ala Val Arg Gly
            115   120   125

Glu Gln Trp Thr Pro Ile Glu Pro Lys Pro Arg Glu Arg Pro Gln Ile
    130   135   140

Page 2

DiffExpr.ST25.txt
```
Gly Gly Thr Ile Lys Gln Pro Pro Ser Asn Pro Pro Pro Arg Pro Pro
145             150             155             160


Ala Glu Ala Arg Lys Lys Pro Ser Glu Glu Glu Ala Ala Ala Ala Ala
            165             170             175


Gly Gly Pro Pro Gly Gly Pro Gln Val Asn Pro Ile Pro Val Thr Asp
            180             185             190


Glu Val Val
        195
```

```
<210>   3
<211>   327
<212>   DNA
<213>   Homo sapiens

<400>   3
aggccaagga tggggtttgg ggtacccccc gaagggaacc taggctcctg tcccggatgg

60

gccgagggga tggaaccgcc cccacaacca gctcctgggg cacggggggag ggatgagaag

120

atgacctgag gggaagctcc tgcagcctcc ttccagcctg ggggctgcat caatgtggct

180

ggtgagccgc gaactgagac gagggaagaa ccgaaccgag gccacccagt aattcccatt

240

gtatcccaaa agaatcttgc caaagatggc gatccggtgc ccgtagcttt gagaagcagt

300

aataacacag gctttgggaa gccagca

327
```

```
<210>   4
<211>   480
<212>   DNA
<213>   Homo sapiens

<400>   4
atgcctgggc accctggggg ggctgccggc cagcgtgggt gaatggcgcc caccagacgg
```

DiffExpr.ST25.txt

60

agcagggcct gctactctcc tgggccaccc ctcctgctgg acctcgggcc tgcctccctg

120

ccttccctgc ccttggcacc ttgagaggag aatttgagta ctagaatttg catctgggag

180

tgtccaccac gggatcttcc agaagcccag gggttcgggc cccactcctg gcatgaagaa

240

atagcttggg ctggtgtcag gccgacccag gtcctggctg tgtggcctcg ggcgcgtcac

300

ttaacctctc tgagtggcag tcacaccatc ctgcgtggga ttcagacttg agcctgtccc

360

agggcccta gcggggtggg cctccgggga gggagggagg gagggagggc gggcgctggg

420

gcgcggctca accattcagt gccccgcagg cggctgtgcg tggcgagcag tggacgccca

480

<210>   5
<211>   890
<212>   DNA
<213>   Homo sapiens

<400>   5
ggcacgaggg cgcaatggga aggagcccac ccacgactgt ggatgtttaa ccaaagcctt

60

cctgtaggaa agcaggaaag actgggtggc ctgagggcaa ggaatggaaa tatcactgaa

120

actaaagaaa acgaccctcc agtccagtgg gtgcaaactg gggacttcag ctggattcac

180

ccctcaggtg gcttctgttt ggccaccacc ttgtttggtc tcccacgtga aaagctggtg

240

cctttcatca cgaagctgct ctcaccccat tcctagattc tgaaacctga agtgggaaat

Page 4

DiffExpr.ST25.txt

300

ccacgttgct ggagaagcca caaaccacac cagccaggtg tgccacccca caccagccag

360

gtgttaggag agagctcaac ctccgtctta gggcctgact tctggtttgc ggctgggtca

420

acaaatccca ctttgcttgg taacaaggag acctgcatcc ctcccaccgg gttcatcttt

480

ccattgacag ttcagcagaa aaggcctttg ggtttgacag ccacccttg agctctccgc

540

agcaggagga aaacgtcaga gagcccttga ggccctgttt cctcaccaga aggaaggtgg

600

ctgacacagc catgatcacc cccaatctgg aggctcagac tagagtggcc ccaaatcctt

660

ggggacaagg ccagaaagtg gatgtggcca cacagacagc tcacagaagt tgaaatctac

720

caaggacatg catggcacac gggggtgtac acacactcgt ttgactcaca cctcatcctc

780

cctttcgctt cagctgggga aggaggtgga aaggtgatgt tcaagactgg agggtaattt

840

gcccaagaat gcagtgagca acatcttaa aacaccacgg agagcctctg

890

<210>   6
<211>   2817
<212>   DNA
<213>   Homo sapiens

<400>   6
ggcacgaggg cgcttcttcg ggtcgcgggt gcactccggc ccggctcccg cctcggcccc

60

gatggacgcc gcgttcctcc tcgtcctcgg gctgttggcc cagagcctct gcctgtcttt

DiffExpr.ST25.txt

120

ggggggttcct ggatggagga ggcccaccac cctgtacccc tggcgccggg cgcctgccct

180

gagccgcgtg cggagggcct gggtcatccc cccgatcagc gtatccgaga accacaagcg

240

tctcccctac cccctggttc agatcaagtc ggacaagcag cagctgggca gcgtcatcta

300

cagcatccag ggacccggcg tggatgagga gccccggggc gtcttctcta tcgacaagtt

360

cacagggaag gtcttcctca atgccatgct ggaccgcgag aagactgatc gcttcaggct

420

aagagcgttt gccctggacc tgggaggatc caccctggag gaccccacgg acctggagat

480

tgtagttgtg gatcagaatg acaaccggcc agccttcctg caggaggcgt tcactggccg

540

cgtgctggag ggtgcagtcc caggcaccta tgtgaccagg gcagaggcca cagatgccga

600

cgaccccgag acggacaacg cagcgctgcg gttctccatc ctgcagcagg cagcccccga

660

gctcttcagc atcgacgagc tcacaggaga gatccgcaca gtgcaagtgg ggctggaccg

720

cgaggtggtc gcggtgtaca atctgaccct gcaggtggcg gacatgtctg agacggcct

780

cacagccact gcctcagcca tcatcaccct tgatgacatc aatgacaatg ccccgagtt

840

caccagggat gagttcttca tggaggccat agaggccgtc agcggagtgg atgtgggacg

900

cctggaagtg gaggacaggg acctgccagg ctccccaaac tgggtggcca ggttcactat

Page 6

DiffExpr.ST25.txt

960

cctggaaggc gaccccgatg ggcagttcac catccgcacg gaccccaaga ccaacgaggg

1020

tgttctgtcc attgtgaagg ccctggacta tgagagctgt gaacactacg aactcaaagt

1080

gtcggtgcag aatgaggccc cgctgcaggc ggctgccctt agggctgagc ggggccaggc

1140

caaggtccgc gtgcatgtgc aggacaccaa cgagcccccc gtgttccagg agaacccact

1200

tcggaccagc ctagcagagg gggcaccccc aggcactctg gtggccacct tctctgcccg

1260

ggaccctgac acagagcagc tgcagaggct cagctactcc aaggactacg acccggaaga

1320

ctggctgcaa gtggacgcag ccactggccg gatccagacc cagcacgtgc tcagcccggc

1380

gtccccttc ctcaagggcg gctggtacag agccatcgtc ctggcccagg atgacgcctc

1440

ccagccccgc accgccaccg gcaccctgtc catcgagatc ctggaggtga cgaccatgc

1500

acctgtgctg gccccgccgc cgccgggcag cctgtgcagc gagccacacc aaggcccagg

1560

cctcctcctg ggcgccacgg atgaggacct gccccccac ggggcccccct tccacttcca

1620

gctgagcccc aggctcccag agctcggccg gaactggagc ctcagccagg tcaacgtgag

1680

ccacgcgcgc ctgcggccgc gacaccaggt ccccgaaggc ctgcaccgcc tcagcctgct

1740

DiffExpr.ST25.txt
gctccgggac tcggggcagc cgccccagca gcgcgagcag cctctgaacg tgaccgtgtg

1800

ccgctgcggc aaggacggcg tctgcctgcc gggggccgca gcgctgctgg cgggggggcac

1860

aggcctcagc ctgggcgcac tggtcatcgt gctggccagc gccctcctgc tgctggtgct

1920

ggtcctgctc gtggcactcc gggcgcggtt ctggaagcag tctcggggca aggggctgct

1980

gcacggcccc caggacgacc ttcgagacaa tgtcctcaac tacgatgagc aaggaggcgg

2040

ggaggaggac caggacgcct acgacatcag ccagctgcgt cacccgacag cgctgagcct

2100

gcctctggga ccgccgccac ttcgcagaga tgccccgcag ggccgcctgc accccagcc

2160

accccgagtg ctgcccacca gccccctgga catcgccgac ttcatcaatg atggcttgga

2220

ggctgcagat agtgacccca gtgtgccgcc ttacgacaca gccctcatct atgactacga

2280

gggtgacggc tcggtggcgg ggacgctgag ctccatcctg tccagccagg gcgatgagga

2340

ccaggactac gactacctca gagactgggg gccccgcttc gcccggctgg cagacatgta

2400

tgggcacccg tgcgggttgg agtacggggc cagatgggac caccaggcca gggagggtct

2460

ttctcctggg gcactgctac ccagacacag aggccggaca gcctgaccct ggggcgcaac

2520

tggacatgcc actccccggc ctcgtggcag tgatggcccc tgcagaggca gcctgaggtc

2580

Page 8

DiffExpr.ST25.txt

accgggcccg accccccctgg gcctggggca gcctccttcc tgtaggcgag ggcccaagtc

2640

tgggggcaga acctgagtgt ggatggggcg gccaggaaga ggccccttcc tgccggggtg

2700

ggaagagttt ctctccatcg gccccatgcg ggtcacctcc ctagtcccac ctttgcctcc

2760

taccagtgaa cctcatcttt gtatgaaaga cagcaacctc ctgggtaaat ctgaatg

2817

<210>   7
<211>   814
<212>   PRT
<213>   Homo sapiens

<400>   7

Met Asp Ala Ala Phe Leu Leu Val Leu Gly Leu Leu Ala Gln Ser Leu
1   5   10   15

Cys Leu Ser Leu Gly Val Pro Gly Trp Arg Arg Pro Thr Thr Leu Tyr
                20   25   30

Pro Trp Arg Arg Ala Pro Ala Leu Ser Arg Val Arg Arg Ala Trp Val
            35   40   45

Ile Pro Pro Ile Ser Val Ser Glu Asn His Lys Arg Leu Pro Tyr Pro
        50   55   60

Leu Val Gln Ile Lys Ser Asp Lys Gln Gln Leu Gly Ser Val Ile Tyr
65   70   75   80

Ser Ile Gln Gly Pro Gly Val Asp Glu Glu Pro Arg Gly Val Phe Ser
                85   90   95

Ile Asp Lys Phe Thr Gly Lys Val Phe Leu Asn Ala Met Leu Asp Arg
                100   105   110

Glu Lys Thr Asp Arg Phe Arg Leu Arg Ala Phe Ala Leu Asp Leu Gly
            115   120   125

Page 9

DiffExpr.ST25.txt

Gly Ser Thr Leu Glu Asp Pro Thr Asp Leu Glu Ile Val Val Val Asp
    130   135   140

Gln Asn Asp Asn Arg Pro Ala Phe Leu Gln Glu Ala Phe Thr Gly Arg
145   150   155   160

Val Leu Glu Gly Ala Val Pro Gly Thr Tyr Val Thr Arg Ala Glu Ala
            165   170   175

Thr Asp Ala Asp Asp Pro Glu Thr Asp Asn Ala Ala Leu Arg Phe Ser
            180   185   190

Ile Leu Gln Gln Gly Ser Pro Glu Leu Phe Ser Ile Asp Glu Leu Thr
        195   200   205

Gly Glu Ile Arg Thr Val Gln Val Gly Leu Asp Arg Glu Val Val Ala
    210   215   220

Val Tyr Asn Leu Thr Leu Gln Val Ala Asp Met Ser Gly Asp Gly Leu
225   230   235   240

Thr Ala Thr Ala Ser Ala Ile Ile Thr Leu Asp Asp Ile Asn Asp Asn
            245   250   255

Ala Pro Glu Phe Thr Arg Asp Glu Phe Phe Met Glu Ala Ile Glu Ala
            260   265   270

Val Ser Gly Val Asp Val Gly Arg Leu Glu Val Glu Asp Arg Asp Leu
        275   280   285

Pro Gly Ser Pro Asn Trp Val Ala Arg Phe Thr Ile Leu Glu Gly Asp
    290   295   300

Pro Asp Gly Gln Phe Thr Ile Arg Thr Asp Pro Lys Thr Asn Glu Gly
305   310   315   320

Val Leu Ser Ile Val Lys Ala Leu Asp Tyr Glu Ser Cys Glu His Tyr
            325   330   335

Glu Leu Lys Val Ser Val Gln Asn Glu Ala Pro Leu Gln Ala Ala Ala
                        Page 10

DiffExpr.ST25.txt
            340   345   350

Leu Arg Ala Glu Arg Gly Gln Ala Lys Val Arg Val His Val Gln Asp
        355   360   365

Thr Asn Glu Pro Pro Val Phe Gln Glu Asn Pro Leu Arg Thr Ser Leu
        370   375   380

Ala Glu Gly Ala Pro Pro Gly Thr Leu Val Ala Thr Phe Ser Ala Arg
385   390   395   400

Asp Pro Asp Thr Glu Gln Leu Gln Arg Leu Ser Tyr Ser Lys Asp Tyr
                405   410   415

Asp Pro Glu Asp Trp Leu Gln Val Asp Ala Ala Thr Gly Arg Ile Gln
            420   425   430

Thr Gln His Val Leu Ser Pro Ala Ser Pro Phe Leu Lys Gly Gly Trp
        435   440   445

Tyr Arg Ala Ile Val Leu Ala Gln Asp Asp Ala Ser Gln Pro Arg Thr
    450   455   460

Ala Thr Gly Thr Leu Ser Ile Glu Ile Leu Glu Val Asn Asp His Ala
465   470   475   480

Pro Val Leu Ala Pro Pro Pro Gly Ser Leu Cys Ser Glu Pro His
                485   490   495

Gln Gly Pro Gly Leu Leu Leu Gly Ala Thr Asp Glu Asp Leu Pro Pro
            500   505   510

His Gly Ala Pro Phe His Phe Gln Leu Ser Pro Arg Leu Pro Glu Leu
        515   520   525

Gly Arg Asn Trp Ser Leu Ser Gln Val Asn Val Ser His Ala Arg Leu
    530   535   540

Arg Pro Arg His Gln Val Pro Glu Gly Leu His Arg Leu Ser Leu Leu
545   550   555   560

Page 11

```
                              DiffExpr.ST25.txt
Leu Arg Asp Ser Gly Gln Pro Pro Gln Gln Arg Glu Gln Pro Leu Asn
            565             570              575


Val Thr Val Cys Arg Cys Gly Lys Asp Gly Val Cys Leu Pro Gly Ala
            580             585              590


Ala Ala Leu Leu Ala Gly Gly Thr Gly Leu Ser Leu Gly Ala Leu Val
            595             600              605


Ile Val Leu Ala Ser Ala Leu Leu Leu Leu Val Leu Val Leu Leu Val
    610             615             620


Ala Leu Arg Ala Arg Phe Trp Lys Gln Ser Arg Gly Lys Gly Leu Leu
625             630             635              640


His Gly Pro Gln Asp Asp Leu Arg Asp Asn Val Leu Asn Tyr Asp Glu
            645             650              655


Gln Gly Gly Gly Glu Glu Asp Gln Asp Ala Tyr Asp Ile Ser Gln Leu
            660             665              670


Arg His Pro Thr Ala Leu Ser Leu Pro Leu Gly Pro Pro Pro Leu Arg
            675             680              685


Arg Asp Ala Pro Gln Gly Arg Leu His Pro Gln Pro Pro Arg Val Leu
    690             695             700


Pro Thr Ser Pro Leu Asp Ile Ala Asp Phe Ile Asn Asp Gly Leu Glu
705             710             715              720


Ala Ala Asp Ser Asp Pro Ser Val Pro Pro Tyr Asp Thr Ala Leu Ile
            725             730              735


Tyr Asp Tyr Glu Gly Asp Gly Ser Val Ala Gly Thr Leu Ser Ser Ile
            740             745              750


Leu Ser Ser Gln Gly Asp Glu Asp Gln Asp Tyr Asp Tyr Leu Arg Asp
            755             760              765


Trp Gly Pro Arg Phe Ala Arg Leu Ala Asp Met Tyr Gly His Pro Cys
    770             775             780
```

Page 12

DiffExpr.ST25.txt

Gly Leu Glu Tyr Gly Ala Arg Trp Asp His Gln Ala Arg Glu Gly Leu
785   790   795   800

Ser Pro Gly Ala Leu Leu Pro Arg His Arg Gly Arg Thr Ala
            805   810

<210> 8
<211> 1566
<212> DNA
<213> Homo sapiens

<400> 8
gtgcctctcc tggcaggcag agtggctcct cacagcctga agctcatcct tctgcacggg

60

ccagccaggc cagcacagag gcaccagggc agcagtgcac acaggtcccc ggggacccca

120

ccatgtggag cggatggtgg ctgtggcccc ttgtggccgt ctgcactgca gacttctttc

180

gggacgaggc agagaggatc atgagggact cccctgtcat tgatgggcac aatgacctcc

240

cctggcagct gctggatatg ttcaacaacc ggctgcagga cgagagggcc aacctgacca

300

ccttggccgg cacacacacc aacatcccca agctgagggc cggctttgtg ggaggccagt

360

tctggtccgt gtacacgccc tgcgacaccc agaacaaaga cgccgtgcgg aggacgctgg

420

agcagatgga cgtggtccac cgcatgtgcc ggatgtaccc ggagaccttc ctgtatgtca

480

ccagcagtgc aggcattcgg caggccttcc gggaagggaa ggtggccagc ctgatcggcg

540

tggagggcgg ccactccatt gacagcagtt tgggcgtcct gcgggcactc tatcagctgg

600

DiffExpr.ST25.txt
gcatgcggta cctgaccctc acccacagct gcaacacgcc ctgggctgac aactggctgg

660

tggacacggg agacagcgag ccccagagcc aaggcttgtc accctttggg cagcgtgtgg

720

tgaaggagct gaaccgtctg ggggtcctca tcgacttggc tcacgtgtct gtggccacca

780

tgaaggccac cctgcagctg tccagagccc cggtcatctt cagccactcc tcggcctaca

840

gcgtgtgcgc aagccggcgc aacgtgcctg acgacgtcct gaggctggtg aaacagacag

900

acagcctggt gatggtgaac ttctacaaca attacatttc ctgcaccaac aaggccaacc

960

tgtcccaagt ggccgaccat ctggatcaca tcaaggaggt ggcaggagcc agagccgtgg

1020

gttttggtgg ggactttgat ggtgttccaa gggtccctga ggggctggag gacgtctcca

1080

agtatccaga cctgatcgct gagctgctca ggaggaactg gacggaggcg gaggtcaagg

1140

gcgcactggc tgacaacctg ctgagggtct tcgaggctgt ggaacaggcc agcaacctca

1200

cacaggctcc cgaggaggag cccatcccgc tggaccagct gggtggctcc tgcaggaccc

1260

attacggcta ctcctctggg gcttccagcc tccatcgcca ctggggggctc ctgctggcct

1320

ccctcgctcc cctggtcctc tgtctgtctc tcctgtgaaa cctgggagac cagagtcccc

1380

tttagggttc ccggagctcc gggaagaccc gcccatccca ggactccaga tgccaggagc

1440

DiffExpr.ST25.txt

cctgctgccc acatgcaagg accagcatct cctgagagga cgcctgggct tacctggggg

1500

gcaggatgcc tggggacagt tcaggacaca cacacagtag gcccgcaata aaagcaacac

1560

cccttc

1566

<210> 9
<211> 411
<212> PRT
<213> Homo sapiens

<400> 9

```
Met Trp Ser Gly Trp Trp Leu Trp Pro Leu Val Ala Val Cys Thr Ala
1               5                   10                  15


Asp Phe Phe Arg Asp Glu Ala Glu Arg Ile Met Arg Asp Ser Pro Val
            20                  25                  30


Ile Asp Gly His Asn Asp Leu Pro Trp Gln Leu Leu Asp Met Phe Asn
            35                  40                  45


Asn Arg Leu Gln Asp Glu Arg Ala Asn Leu Thr Thr Leu Ala Gly Thr
        50                  55                  60


His Thr Asn Ile Pro Lys Leu Arg Ala Gly Phe Val Gly Gly Gln Phe
65                  70                  75                  80


Trp Ser Val Tyr Thr Pro Cys Asp Thr Gln Asn Lys Asp Ala Val Arg
                85                  90                  95


Arg Thr Leu Glu Gln Met Asp Val Val His Arg Met Cys Arg Met Tyr
            100                 105                 110


Pro Glu Thr Phe Leu Tyr Val Thr Ser Ser Ala Gly Ile Arg Gln Ala
            115                 120                 125


Phe Arg Glu Gly Lys Val Ala Ser Leu Ile Gly Val Glu Gly Gly His
    130                 135                 140
```

Page 15

DiffExpr.ST25.txt

Ser Ile Asp Ser Ser Leu Gly Val Leu Arg Ala Leu Tyr Gln Leu Gly
145   150   155   160

Met Arg Tyr Leu Thr Leu Thr His Ser Cys Asn Thr Pro Trp Ala Asp
                 165   170   175

Asn Trp Leu Val Asp Thr Gly Asp Ser Glu Pro Gln Ser Gln Gly Leu
             180   185   190

Ser Pro Phe Gly Gln Arg Val Val Lys Glu Leu Asn Arg Leu Gly Val
           195   200   205

Leu Ile Asp Leu Ala His Val Ser Val Ala Thr Met Lys Ala Thr Leu
     210   215   220

Gln Leu Ser Arg Ala Pro Val Ile Phe Ser His Ser Ser Ala Tyr Ser
225   230   235   240

Val Cys Ala Ser Arg Arg Asn Val Pro Asp Asp Val Leu Arg Leu Val
             245   250   255

Lys Gln Thr Asp Ser Leu Val Met Val Asn Phe Tyr Asn Asn Tyr Ile
             260   265   270

Ser Cys Thr Asn Lys Ala Asn Leu Ser Gln Val Ala Asp His Leu Asp
             275   280   285

His Ile Lys Glu Val Ala Gly Ala Arg Ala Val Gly Phe Gly Gly Asp
     290   295   300

Phe Asp Gly Val Pro Arg Val Pro Glu Gly Leu Glu Asp Val Ser Lys
305   310   315   320

Tyr Pro Asp Leu Ile Ala Glu Leu Leu Arg Arg Asn Trp Thr Glu Ala
             325   330   335

Glu Val Lys Gly Ala Leu Ala Asp Asn Leu Leu Arg Val Phe Glu Ala
           340   345   350

Val Glu Gln Ala Ser Asn Leu Thr Gln Ala Pro Glu Glu Glu Pro Ile

DiffExpr.ST25.txt
```
            355                    360                    365
```

```
Pro Leu Asp Gln Leu Gly Gly Ser Cys Arg Thr His Tyr Gly Tyr Ser
    370                 375                 380


Ser Gly Ala Ser Ser Leu His Arg His Trp Gly Leu Leu Leu Ala Ser
385                 390                 395                 400


Leu Ala Pro Leu Val Leu Cys Leu Ser Leu Leu
              405             410
```

```
<210>  10
<211>  4539
<212>  DNA
<213>  Homo sapiens

<400>  10
cggcgcgcac actgctcgct gggccgcggc tcccgggtgt cccaggcccg gccggtgcgc
60

agagcatggc gggtgcgggc ccgaagcggc gcgcgctagc ggcgccggcg ccgaggaga
120

aggaagaggc gcgggagaag atgctggccg ccaagagcgc ggacggctcg cgccggcag
180

gcgagggcga gggcgtgacc ctgcagcgga acatcacgct gctcaacggc gtggccatca
240

tcgtggggac cattatcggc tcgggcatct tcgtgacgcc cacgggcgtg ctcaaggagg
300

caggctcgcc ggggctggcg ctggtggtgt gggccgcgtg cggcgtcttc tccatcgtgg
360

gcgcgctctg ctacgcggag ctcggcacca ccatctccaa atcgggcggc gactacgcct
420

acatgctgga ggtctacggc tcgctgcccg ccttcctcaa gctctggatc gagctgctca
480

tcatccggcc ttcatcgcag tacatcgtgg ccctggtctt cgccacctac ctgctcaagc
```

DiffExpr.ST25.txt

540

cgctcttccc cacctgcccg gtgcccgagg aggcagccaa gctcgtggcc tgcctctgcg

600

tgctgctgct cacggccgtg aactgctaca gcgtgaaggc cgccacccgg gtccaggatg

660

cctttgccgc cgccaagctc ctggccctgg ccctgatcat cctgctgggc ttcgtccaga

720

tcgggaaggg tgatgtgtcc aatctagatc ccaacttctc atttgaaggc accaaactgg

780

atgtggggaa cattgtgctg gcattataca gcggcctctt tgcctatgga ggatggaatt

840

acttgaattt cgtcacagag gaaatgatca acccctacag aaacctgccc ctggccatca

900

tcatctccct gcccatcgtg acgctggtgt acgtgctgac caacctggcc tacttcacca

960

ccctgtccac cgagcagatg ctgtcgtccg aggccgtggc cgtggacttc gggaactatc

1020

acctgggcgt catgtcctgg atcatccccg tcttcgtggg cctgtcctgc ttcggctccg

1080

tcaatgggtc cctgttcaca tcctccaggc tcttcttcgt ggggtcccgg gaaggccacc

1140

tgccctccat cctctccatg atccacccac agctcctcac ccccgtgccg tccctcgtgt

1200

tcacgtgtgt gatgacgctg ctctacgcct tctccaagga catcttctcc gtcatcaact

1260

tcttcagctt cttcaactgg ctctgcgtgg ccctggccat catcggcatg atctggctgc

1320

gccacagaaa gcctgagctt gagcggccca tcaaggtgaa cctggccctg cctgtgttct

DiffExpr.ST25.txt

1380

tcatcctggc ctgcctcttc ctgatcgccg tctccttctg gaagacaccc gtggagtgtg

1440

gcatcggctt caccatcatc ctcagcgggc tgcccgtcta cttcttcggg gtctggtgga

1500

aaaacaagcc caagtggctc ctccagggca tcttctccac gaccgtcctg tgtcagaagc

1560

tcatgcaggt ggtcccccag gagacatagc caggaggccg agtggctgcc ggaggagcat

1620

gcgcagaggc cagttaaagt agatcacctc ctcgaaccca ctccggttcc ccgcaaccca

1680

cagctcagct gcccatccca gtccctcgcc gtccctccca ggtcgggcag tggaggctgc

1740

tgtgaaaact ctggtacgaa tctcatccct caactgaggg ccagggaccc aggtgtgcct

1800

gtgctcctgc ccaggagcag cttttggtct ccttgggccc tttttccctt ccctcctttg

1860

tttacttata tatatatttt ttttaaactt aaattttggg tcaacttgac accactaaga

1920

tgattttta aggagctggg ggaaggcagg agccttcctt tctcctgccc caagggccca

1980

gaccctgggc aaacagagct actgagactt ggaacctcat tgctacgaca gacttgcact

2040

gaagccggac agctgcccag acacatgggc ttgtgacatt cgtgaaaacc aaccctgtgg

2100

gcttatgtct ctgccttagg gtttgcagag tggaaactca gccgtagggt ggcactggga

2160

Page 19

DiffExpr.ST25.txt

ggggggtgggg gatctgggca aggtgggtga ttcctcccag gaggtgcttg aggccccgat

2220

ggactcctga ccataatcct agccccgaga caccatcctg agccagggaa cagccccagg

2280

gttggggggt gccggcatct cccctagctc accaggcctg gcctctgggc agtgtggcct

2340

cttggctatt tctgttccag ttttggaggc tgagttctgg ttcatgcaga caaagccctg

2400

tccttcagtc ttctagaaac agagacaaga aaggcagaca caccgcggcc aggcacccat

2460

gtgggcgccc accctgggct ccacacagca gtgtcccctg ccccagaggt cgcagctacc

2520

ctcagcctcc aatgcattgg cctctgtacc gcccggcagc cccttctggc cggtgctggg

2580

ttcccactcc cggcctaggc acctccccgc tctccctgtc acgctcatgt cctgtcctgg

2640

tcctgatgcc cgttgtctag gagacagagc caagcactgc tcacgtctct gccgcctgcg

2700

tttggaggcc cctgggctct cacccagtcc ccacccgcct gcagagaggg aactagggca

2760

ccccttgttt ctgttgttcc cgtgaatttt tttcgctatg ggaggcagcc gaggcctggc

2820

caatgcggcc cactttcctg agctgtcgct gcctccatgg cagcagccaa ggacccccag

2880

aacaagaaga ccccccccgca ggatccctcc tgagctcggg gggctctgcc ttctcaggcc

2940

ccgggcttcc cttctcccca gccagaggtg gagccaagtg gtccagcgtc actccagtgc

3000

DiffExpr.ST25.txt

tcagctgtgg ctggaggagc tggcctgtgg cacagccctg agtgtcccaa gccgggagcc

3060

aacgaagccg gacacggctt cactgaccag cggctgctca agccgcaagc tctcagcaag

3120

tgcccagtgg agcctgccgc ccccacctgg gcaccgggac cccctcacca tccagtgggc

3180

ccggagaaac ctgatgaaca gtttggggac tcaggaccag atgtccgtct ctcttgcttg

3240

aggaatgaag acctttattc acccctgccc cgttgcttcc cgctgcacat ggacagactt

3300

cacagcgtct gctcatagga cctgcatcct tcctggggac gaattccact cgtccaaggg

3360

acagcccacg gtctggaggc cgaggaccac cagcaggcag gtggactgac tgtgttgggc

3420

aagacctctt ccctctgggc ctgttctctt ggctgcaaat aaggacagca gctggtgccc

3480

cacctgcctg gtgcattgct gtgtgaatcc aggaggcagt ggacatcgta ggcagccacg

3540

gccccgggtc caggagaagt gctccctgga ggcacgcacc actgcttccc actggggccg

3600

gcggggccca cgcacgacgt cagcctctta ccttcccgcc tcggctaggg gtcctcggga

3660

tgccgttctg ttccaacctc ctgctctggg aggtggacat gcctcaagga tacagggagc

3720

cggcggcctc tcgacggcac gcacttgcct gttggctgct gcggctgtgg gcgagcatgg

3780

gggctgccag cgtctgttgt ggaaagtagc tgctagtgaa atggctgggg ccgctggggt

DiffExpr.ST25.txt

3840

ccgtcttcac actgcgcagg tctcttctgg gcgtctgagc tggggtggga gctcctccgc

3900

agaaggttgg tgggggggtcc agtctgtgat ccttggtgct gtgtgcccca ctccagcctg

3960

gggaccccac ttcagaaggt aggggccgtg tcccgcggtg ctgactgagg cctgcttccc

4020

cctcccctc ctgctgtgct ggaattccac agggaccagg gccaccgcag gggactgtct

4080

cagaagactt gattttccg tccctttttc tccacactcc actgacaaac gtccccagcg

4140

gtttccactt gtgggcttca ggtgttttca agcacaaccc accacaacaa gcaagtgcat

4200

tttcagtcgt tgtgctttt tgttttgtgc taacgtctta ctaatttaaa gatgctgtcg

4260

gcaccatgtt tatttatttc cagtggtcat gctcagcctt gctgctctgc gtggcgcagg

4320

tgccatgcct gctccctgtc tgtgtcccag ccacgcaggg ccatccactg tgacgtcggc

4380

cgaccaggct ggacaccctc tgccgagtaa tgacgtgtgt ggctgggacc ttctttattc

4440

tgtgttaatg gctaacctgt tacactgggc tgggttgggt agggtgttct ggcttttttg

4500

tggggttttt atttttaaag aaacactcaa tcatcctag

4539

<210> 11
<211> 507
<212> PRT
<213> Homo sapiens

DiffExpr.ST25.txt

<400> 11

Met Ala Gly Ala Gly Pro Lys Arg Arg Ala Leu Ala Ala Pro Ala Ala
1   5   10   15

Glu Glu Lys Glu Glu Ala Arg Glu Lys Met Leu Ala Ala Lys Ser Ala
                20   25   30

Asp Gly Ser Ala Pro Ala Gly Glu Gly Glu Gly Val Thr Leu Gln Arg
            35   40   45

Asn Ile Thr Leu Leu Asn Gly Val Ala Ile Ile Val Gly Thr Ile Ile
        50   55   60

Gly Ser Gly Ile Phe Val Thr Pro Thr Gly Val Leu Lys Glu Ala Gly
65   70   75   80

Ser Pro Gly Leu Ala Leu Val Val Trp Ala Ala Cys Gly Val Phe Ser
                85   90   95

Ile Val Gly Ala Leu Cys Tyr Ala Glu Leu Gly Thr Thr Ile Ser Lys
            100   105   110

Ser Gly Gly Asp Tyr Ala Tyr Met Leu Glu Val Tyr Gly Ser Leu Pro
            115   120   125

Ala Phe Leu Lys Leu Trp Ile Glu Leu Leu Ile Ile Arg Pro Ser Ser
    130   135   140

Gln Tyr Ile Val Ala Leu Val Phe Ala Thr Tyr Leu Leu Lys Pro Leu
145   150   155   160

Phe Pro Thr Cys Pro Val Pro Glu Glu Ala Ala Lys Leu Val Ala Cys
            165   170   175

Leu Cys Val Leu Leu Leu Thr Ala Val Asn Cys Tyr Ser Val Lys Ala
            180   185   190

Ala Thr Arg Val Gln Asp Ala Phe Ala Ala Ala Lys Leu Leu Ala Leu
        195   200   205

Page 23

DiffExpr.ST25.txt

```
Ala Leu Ile Ile Leu Leu Gly Phe Val Gln Ile Gly Lys Gly Asp Val
    210             215             220

Ser Asn Leu Asp Pro Asn Phe Ser Phe Glu Gly Thr Lys Leu Asp Val
225             230             235                         240

Gly Asn Ile Val Leu Ala Leu Tyr Ser Gly Leu Phe Ala Tyr Gly Gly
                245             250             255

Trp Asn Tyr Leu Asn Phe Val Thr Glu Glu Met Ile Asn Pro Tyr Arg
            260             265             270

Asn Leu Pro Leu Ala Ile Ile Ile Ser Leu Pro Ile Val Thr Leu Val
        275             280             285

Tyr Val Leu Thr Asn Leu Ala Tyr Phe Thr Thr Leu Ser Thr Glu Gln
    290             295             300

Met Leu Ser Ser Glu Ala Val Ala Val Asp Phe Gly Asn Tyr His Leu
305             310             315                         320

Gly Val Met Ser Trp Ile Ile Pro Val Phe Val Gly Leu Ser Cys Phe
            325             330             335

Gly Ser Val Asn Gly Ser Leu Phe Thr Ser Ser Arg Leu Phe Phe Val
            340             345             350

Gly Ser Arg Glu Gly His Leu Pro Ser Ile Leu Ser Met Ile His Pro
        355             360             365

Gln Leu Leu Thr Pro Val Pro Ser Leu Val Phe Thr Cys Val Met Thr
    370             375             380

Leu Leu Tyr Ala Phe Ser Lys Asp Ile Phe Ser Val Ile Asn Phe Phe
385             390             395                         400

Ser Phe Phe Asn Trp Leu Cys Val Ala Leu Ala Ile Ile Gly Met Ile
            405             410             415

Trp Leu Arg His Arg Lys Pro Glu Leu Glu Arg Pro Ile Lys Val Asn
        420             425             430
```

DiffExpr.ST25.txt

Leu Ala Leu Pro Val Phe Phe Ile Leu Ala Cys Leu Phe Leu Ile Ala
        435   440   445

Val Ser Phe Trp Lys Thr Pro Val Glu Cys Gly Ile Gly Phe Thr Ile
        450   455   460

Ile Leu Ser Gly Leu Pro Val Tyr Phe Phe Gly Val Trp Trp Lys Asn
465   470   475   480

Lys Pro Lys Trp Leu Leu Gln Gly Ile Phe Ser Thr Thr Val Leu Cys
            485   490   495

Gln Lys Leu Met Gln Val Val Pro Gln Glu Thr
            500   505

Page 25

A.	CLASSIFICATION OF SUBJECT MATTER

Int. Cl. [7]:   C12N 15/63 15/12, A61K 39/395 48/00, C07K 16/18 14/435

According to International Patent Classification (IPC) or to both national classification and IPC

B.	FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

SEE ELECTRONIC DATABASE BELOW

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SEE ELECTRONIC DATABASE BOX BELOW

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EMBL, SWISS-PROT, PIR, EMBL, GENEBANK: SEQ ID Nos: 1-11

C.	DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GenPept database Accession Number: AAA90925. *Parkos, C.A et al.* human neutrophil cytochrome b light chain p22 phagocyte b-cytochrome mRNA. 30 June 1988. & Swiss-Prot database Accession Number C24A Human. 1 January 1990. See whole document. Amino acid sequence is 100% identical to SEQ ID No: 2 (CYBA) over 195 & 194 amino acids respectively.	12, 14, 16, 33, 42 & 54

[X] Further documents are listed in the continuation of Box C	[ ] See patent family annex

*	Special categories of cited documents:
"A"	Document defining the general state of the art which is not considered to be of particular relevance
"E"	Earlier application or patent but published on or after the international filing date
"L"	Document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O"	Document referring to an oral disclosure, use, exhibition or other means
"P"	Document published prior to the international filing date but later than the priority date claimed

"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 March 2002	- 8 APR 2002

Name and mailing address of the ISA/AU	Authorized officer
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	PHILIPPA WYRDEMAN Telephone No : (02) 6283 2554

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Genomics (1999) 60(3): 295-308. *Loftus, B.J et al.* "Genome duplications and other features in 12 mb of DNA sequence from human chromosome 16p and 16q". & EMBL database Accession No: AC003007. 29 October 1997. See whole document. Nucleic acid sequence is 100% identical to SEQ ID No: 3 over 327 nucleic acids.	12, 14, 16, 18-32, 42, 54, 66 & 67
X	EMBL database Accession No: AC005632. *Sulston, J.E et al.* Homo sapiens , BAC clone CTB-31N19 from 16. 9 September 1998. See whole document. Nucleic acid sequence is 100% identical to SEQ ID No: 3 over 327 nucleic acids.	12, 14, 16, 18-32, 42, 54, 66 & 67
X	EMBL database Accession No: AC009086. DOE Joint Genome Institute. Homo sapiens chromosome 16 clone RP11-368N21. 4 August 1999. See whole document. Nucleic acid sequence is 99% identical to SEQ ID No: 3 over 327 nucleic acids.	12, 14, 16, 18-32, 42, 54, 66 & 67
X	J. Biol. Chem. (1998) 273: 10011-10018. *Shimoyama, Y et al.* "Molecular cloning and characterisation of a novel human classic cadherin with mouse mucle cadherin". & EMBL database Accession No: D83542, 29 February 1996. See whole document. Nucleic acid sequence is 99.6% identical to SEQ ID No: 6 (CDH15) over 2817 nucleic acids.	1, 2, 4-7, 12, 14, 16, 33, 54, 66 & 67
X	Biochem Biophys Res Commun (1999) 255(2):283-8. *Prasad, P.D et al.* "Human LAT1, a subunit of system L amino acid transporter: molecular cloning and transport function". & EMBL database Accession No: AF104032, 2 November 1998. See whole document. Nucleic acid sequence is 99% identical to SEQ ID No: 10 (SLC7A5) over 4539 nucleic acids.	13, 15, 17 & 33
X	EMBL database Accession No: AB018009. Cloning and characterisation of a human system transporter. 26 September 1998. See whole document. Nucleic acid sequence is 100% identical to SEQ ID No: 10 (SLC7A5) over 4539 nucleic acids.	1, 3, 8-11, 35-42, 51, 53, 55, 56, 58, 59 & 67-69
X	Nature (1998) 395(6699): 288-291. *Mastroberardino, L et at.* "Amino acid transport by heterodimers of 4F2hc/CD98 and members of permease family". See whole document. Amino acid sequence is 99% identical to SEQ ID No: 11 (SLC7A5) over 507 amino acids. & EMBL database Accession Nos: Q01650 & AAC61479.	13, 15, 17 & 33
X	EMBL database Accession No: BAA75746. *Minato, N et al.* 4F2 light chain [Homo sapiens] 9 March 1999. See whole document. Amino acid sequence is 98% identical to SEQ ID No: 11 (SLC7A5) over 507 amino acids.	13, 15, 17 & 33

Box I	**Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos :

   because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos : **60-65**

   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

   See supplemental box

3. ☐ Claims Nos :

   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II	**Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest   ☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Supplemental Box
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: 1

The scope of claims 60-65 is so inadequately supported by the specification that no meaningful search could be carried out.

These claims refer to the use of an anti-cancer drug or an anti-cancer drug when identified by the method of the invention or a compound for the increased activity/expression of a gene or an antagonist of a gene. However the claims are not restricted to "compounds" (including drugs and antagonists) produced using the applicant invention because the process of identifying "compounds" does not produce a product, it simply further characterises pre-existing "compounds". As such, the claims include within their scope "compounds", including pre-existing known compounds, whose engineering or isolation owes nothing to the teachings if the patent application. Therefore the only association between the applicants invention and the compounds is that the invention has been used to further characterise the inherent properties of pre-existing compounds.

o